CHARACTERIZING THE DOMAINS OF THE BIOFILM REGULATOR MBAA IN *VIBRIO CHOLERAE*

A Thesis
by
WHITNEY GABRIELLE BOND

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTERS OF SCIENCE

August 2018
Department of Biology

CHARACTERIZING THE DOMAINS OF THE BIOFILM REGULATOR MBAA IN *VIBRIO CHOLERAE*

A Thesis
by
WHITNEY GABRIELLE BOND
August 2018

APPROVED BY:

Dr. Ece Karatan
Chairperson, Thesis Committee

Dr. Andrew Bellemer
Member, Thesis Committee

Dr. Darren Seals
Member, Thesis Committee

Dr. Zack Murrell
Chairperson, Department of Biology

Dr. Michael McKenzie,
Dean, Cratis D. Williams School of Graduate Studies

Abstract

CHARACTERIZING THE DOMAINS OF THE BIOFILM REGULATOR MBAA IN *VIBRIO CHOLERAE*

Whitney Gabrielle Bond
B.Sc., University of North Carolina Wilmington
M.A., Appalachian State University

Chairperson: Dr. Ece Karatan

Vibrio cholerae is an aquatic, intestinal pathogen that is believed to survive within its natural environment in a biofilm state but can only cause infection in a planktonic, motile state. This bacterium is able to switch between states by sensing signals within its environment. Polyamines, specifically norspermidine, spermidine, and spermine, modulate biofilm formation. Spermine and spermidine decrease biofilm formation and are found in the human intestine, while norspermidine increase biofilm formation and is found in the environment. All three polyamines are sensed by the periplasmic protein NspS, when NspS binds to a polyamine it is hypothesized to bind to the periplasmic region of MbaA, which is able to transduce a signal into the cytoplasm. MbaA is an integral membrane protein that has C-terminal tandem GGDEF and EAL domains. GGDEF and EAL domains are associated with diguanylate cyclase and phosphodiesterase activity, respectively. Diguanylate cyclases produce the secondary messenger c-di-GMP from GTP while phosphodiesterases break down c-di-GMP. Increasing intracellular c-di-GMP enhances biofilm formation. Our lab has previously characterized MbaA as a phosphodiesterase. This current study set out to

determine the purpose of the GGDEF domain within MbaA. The parts of the *mbaA* gene that encode the GGDEF and EAL domains were amplified by PCR and then cloned into pMAL-c5x. This enabled the expression of GGDEF and EAL domains as maltose binding protein (MBP) fusions. MBP-GGDEF and MBP-EAL proteins were purified and used to determine diguanylate cyclase and phosphodiesterase activities. High Pressure Liquid Chromatography (HPLC) was then used to determine reaction products. The EAL domain was determined to be independently functional *in vitro* and the GGDEF domain has no apparent activity. Chromosomal deletions of each domain were done to further characterize the domains. EAL and GGDEF domain deletions have similar phenotypes to the *mbaA* background. This indicates that both domains are necessary for phosphodiesterase function *in vivo*. The domains were then deleted within an overexpression plasmid pMbaA to further characterize its effect on biofilm formation. pMbaAΔGGDEF had a decreased biofilm formation similar to that of the biofilm formation of pMbaA. pMbaAΔEAL had an increased biofilm formation similar to that of a *ΔmbaA* mutant. This indicates that the GGDEF domain is not necessary for phosphodiesterase activity and confirms *in vitro* results.

Keywords *Vibrio cholerae*; biofilm; phosphodiesterase; diguanylate cyclase; c-di-GMP turnover; GGDEF-EAL protein.

Acknowledgments

I would like to thank the Department of Biology for granting me a teaching assistantship. I would like to acknowledge the Office of Student Research and the Wayne Richardson Fellowship for its support of this project in the form of monetary grants for both research and travel. This project was also supported in part by NIH grant AI096358 from the National Institute of Allergy and Infectious Diseases awarded to Dr. Ece Karatan. I want to thank my committee members Dr. Andrew Bellemer and Dr. Darren Seals for their time and help throughout my research. I would like to also thank Dr. Ece Karatan, my thesis advisor, for her guidance, patience, and brain power.

Dedication

I dedicate this thesis to my husband and dog, who have been my rocks throughout graduate school.

Table of Contents

List of Tables

List of Figures

Introduction

Vibrio cholerae as a pathogen

Vibrio cholerae is a Gram-negative bacterium that is the cause of the human intestinal disease cholera. Infections begin with the ingestion of contaminated water and foodstuffs and progresses to a watery diarrhea (Faruque *et al.*, 1998, Reidl & Klose, 2002). The body fluid loss can be as much as 20 L a day in adults and can lead to death in as few as 12 hours after first symptoms occur (Faruque *et al.*, 1998). Cholera infections were first documented in India in 1817 and spread throughout the world with infections reported in Europe and North America in the early 1900s (Faruque *et al.*, 1998, Reidl & Klose, 2002). Modern sanitation systems prevent cholera outbreaks from happening today in developed countries; however, developing countries such as India, Bangladesh, and parts of Africa still see deadly outbreaks (Faruque *et al.*, 1998). The World Health Organization estimates 1.3 million to 4.0 million cases of cholera happen worldwide each year with 21,000 to 143,000 of those infections ending in death. It is believed that there could be 100 times more cases that go unreported (Faruque *et al.*, 1998, WHO, 2018). The deadliness of this bacteria and how it persists within the environment are research topics still being examined today.

V. cholerae is found in many riverine, estuarine, and marine environments across the world and it is estimated to become more abundant as temperatures continue to warm (Kierek & Watnick, 2003, Escobar *et al.*, 2015). It is thought that *V. cholerae* is able to persist within the environment through its ability to form a community-based, sedentary state called a

biofilm (Reidl & Klose, 2002, Kierek & Watnick, 2003). Biofilms have been shown to provide bacteria with protection from predators, osmotic stress, UV radiation, extreme pH, and the host immune system (Elasri & Miller, 1999, Prigent-Combaret *et al.*, 1999, Donlan & Costerton, 2002, McNeill & Hamilton, 2003). This biofilm state is not only hypothesized to help survival of *V. cholerae* in the environment but also to facilitate the beginning of cholera infections. In order to start an infection, *V. cholerae* has to pass through the gastric barrier to make it into the small intestine. It is thought that the biofilm's protective nature aids bacteria in surviving through the acidic pH of the stomach because *V. cholerae* in biofilms are 1,000-fold more resistant to low pH compared to bacteria in a planktonic state (Zhu & Mekalanos, 2003). Once the *V. cholerae* progress into the small intestine the bacteria need to disperse out of the biofilm and become motile and solitary planktonic cells in order to start expressing the essential virulence factors: cholerae toxin (CT) and toxin-coregulated pilus (TCP) (Reidl & Klose, 2002, Teschler *et al.*, 2015). TCP is required for colonization and infection. It is believed to bind to the intestinal epithelial cells and play a part in microcolony formation (Silva & Benitez, 2016). CT is an A-B type protein toxin that causes the characteristic profuse watery diarrhea seen in cholera infections. The B subunit binds to the GM1 ganglioside receptors in the intestine (Holmgren *et al.*, 1975, Vanden Broeck *et al.*, 2007). The whole toxin is then absorbed into the cell where several CT proteins will cluster in the endoplasmic reticulum (ER); when it is in the ER, the A subunit splits from the B subunit (Holmgren *et al.*, 1975, Vanden Broeck *et al.*, 2007, Sanchez & Holmgren, 2011). This dissociation activates the enzymatic activity of the A subunit. The A subunit then moves back to the cytosol and uses NAD^+ to ADP ribosylate the α subunit of a stimulatory G protein (Gsα) that regulates the activity of adenylate cyclase. ADP ribosylation of Gsα results in

constitutive activation adenylate cyclase which increases cyclic AMP levels in the host cell. This results in protein kinase A phosphorylating a chloride ion channel causing an increase in secretion of chloride ions followed by secretion of water out of the cells, leading to the watery diarrhea seen in cholera infections (Sanchez & Holmgren, 2011). As the disease progresses, *V. cholerae* is passed in the stools of patients (Reidl & Klose, 2002, Teschler *et al.*, 2015). The idea that *V. cholerae* must be in the biofilm to survive stressors but planktonic to cause disease suggests that the switch between biofilm and planktonic state must be highly regulated (Teschler *et al.*, 2015).

C-di-GMP regulation within *V. cholerae*

In order to regulate this switch between biofilm and planktonic states, *V. cholerae* uses environmental stimuli such as bile salts, temperatures, and polyamines (Conner *et al.*, 2017). *V. cholerae* senses these stimuli and then modulates the internal concentration of the second messenger cyclic diguanylate monophosphate (c-di-GMP). High c-di-GMP levels enhance biofilm formation and decrease the expression of both motility and virulence factors. C-di-GMP controls these phenotypes through its effects on transcription, translation, and protein function (Conner *et al.*, 2017).

C-di-GMP controls biofilm formation from the very beginning by controlling the attachment to a surface. *V. cholerae* uses the type IV pilus, MshA, to sense and then bind to appropriate surfaces to begin biofilm formation (Teschler *et al.*, 2015, Conner *et al.*, 2017). This pilus is formed by polymerization of MshA subunits with the energy being provided by the AAA+ ATPase MshE. The MshE domain, MshEN, binds to c-di-GMP and activates the polymerization of MshA subunits (Fig. 1a) (Wang *et al.*, 2016). After the cells have attached themselves to a favorable surface, they begin producing biofilm matrix components such as

Vibrio polysaccharide (VPS) and extracellular proteins. This matrix joins cells together and is a necessary part of the biofilm. The genes that encode the extracellular proteins and the VPS synthesizing proteins are controlled by the biofilm master regulator VpsR (Fig. 1 B). While VpsR can bind to c-di-GMP, this binding is not critical for DNA binding which is required to activate transcription. The significance of VpsR binding to c-di-GMP has yet to be elucidated (Teschler *et al.*, 2015, Conner *et al.*, 2017). VpsR also is necessary for the expression of the accessory biofilm transcriptional regulator VpsT (Srivastava *et al.*, 2011). VpsT then increases the transcription of *vpsR*, *vpsT*, *vps* genes, and genes related to extracellular proteins, such as *rbmA* and *rbmB* (Fig. 1b). VpsT is a homodimer and this dimerization is thought to be stabilized by the binding of 2 intertwined c-di-GMP molecules (Conner *et al.*, 2017). The binding of c-di-GMP to VpsT is necessary for transcription of biofilm genes (Teschler *et al.*, 2015, Conner *et al.*, 2017).

While c-di-GMP enhances biofilm formation, it represses motility. *V. cholerae* is able to move through its environment by a single polar flagellum. C-di-GMP controls flagellar activity both at the transcriptional and functional levels (Conner *et al.*, 2017). The synthesis of the flagellum requires a multitude of proteins. The genes that encode these proteins are regulated by a four-level transcriptional hierarchy. The transcription of flagellar genes is initiated by the transcriptional activator FlrA. C-di-GMP binds to FlrA, inhibiting it from activating the *flrBC* operon, which in turn, inhibits flagellar production (Fig. 1c) (Srivastava *et al.*, 2011).

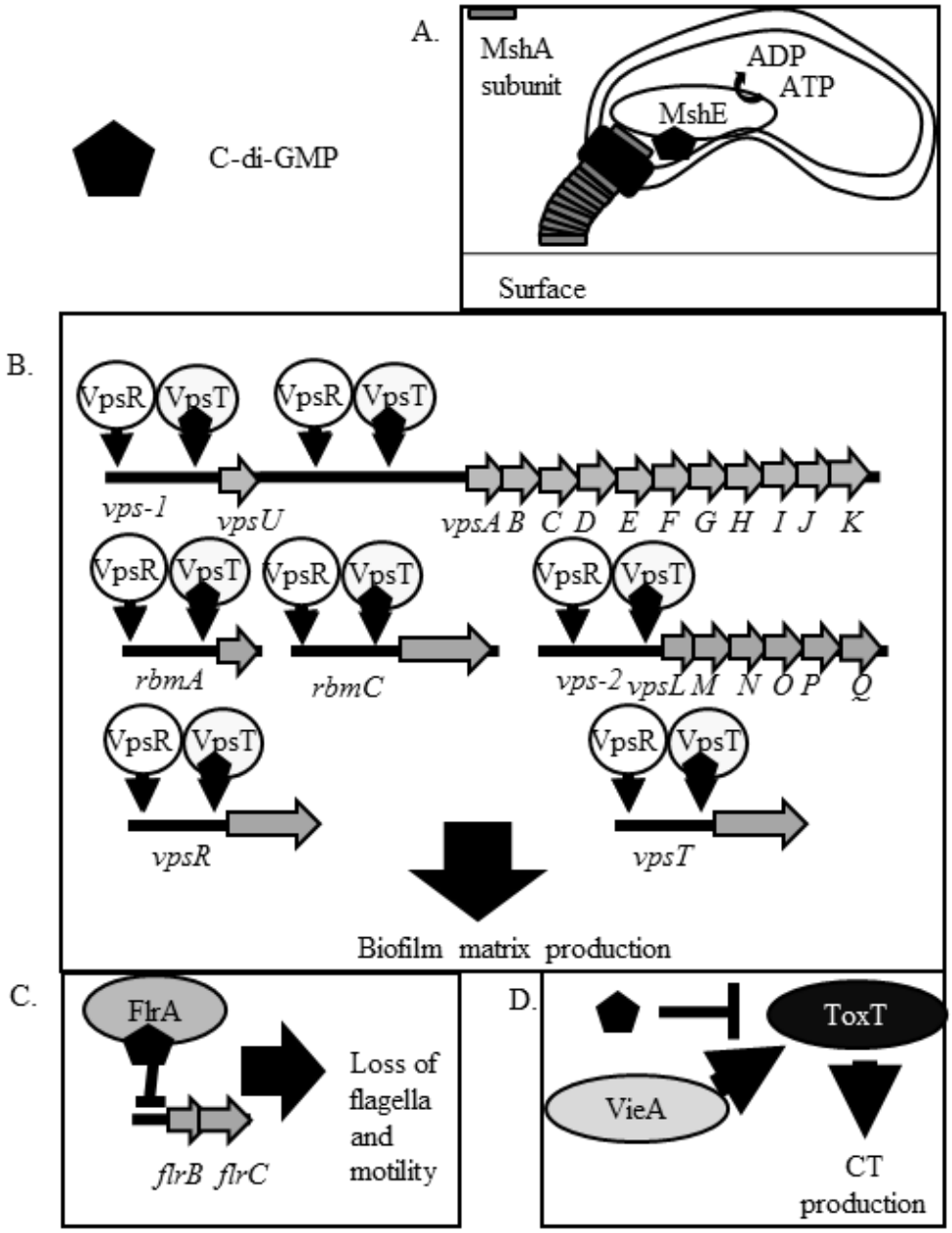

Fig 1. c-di-GMP control of *Vibrio cholerae*. (a) c-di-GMP **(⬟)** binds to the ATPase MshE which polymerizes the MshA subunits (dark gray bars). MshA is a pili necessary for surface attachment and biofilm formation. (b) VpsR and VpsT are transcriptional regulators that bind to promoter regions of biofilm genes such as *vps-1* and *vps-2* operons, *rbmA* and *rbmC*, and *vpsT* and *vpsR*. This leads to biofilm matrix production (c) c-di-GMP binds to the transcriptional regulator FlrA which prevents the flagellar genes *flrB* and *flrC* from being transcribed and causes the inhibition of motility. (d) VieA breaks c-di-GMP down which leads to the production of ToxT. This leads to production of cholera toxin (CT).

High c-di-GMP decreases the transcription of the master virulence transcription factor ToxT. ToxT activates the necessary virulence factors CT and TCP. The loss of a protein called VieA, which breaks down c-di-GMP, leads to an increase in c-di-GMP levels, and this leads to a decrease in ToxT levels and loss of colonization (Tischler & Camilli, 2005). VieA transcription is upregulated when cells adhered to intestinal epithelial cells which leads to a decrease in c-di-GMP and an increase in virulence gene expression (Dey *et al.*, 2013).

Synthesis and breakdown of c-di-GMP

C-di-GMP levels in the cell are controlled by two classes of proteins: diguanylate cyclases and phosphodiesterases. Diguanylate cyclases circularize two guanosine tri-phosphates to produce c-di-GMP; this enzyme is associated with a GGDEF (also called a GG(E/D)EF) domain. Phosphodiesterases are associated with the domains EAL and HD-GYP; these enzymes break down c-di-GMP into linear 5'-phosphoguanylyl-(3',5')-guanosine (5’pGpG) and 2 guanosine monophosphate (GMP), respectively. These two classes of proteins control c-di-GMP concentration within the cell, regulating this switch between planktonic and biofilm states (Conner *et al.*, 2017).

The catalytic amino acids necessary for diguanylate cyclase activity were first characterized through the crystallization of the *Caulobacter crescentus* diguanylate cyclase PleD. GTP binds to the domain between two α helices with the ribosyl and α phosphates covering the glycines within the characteristic GGEEF motif (Chan *et al.*, 2004). Glycines are the smallest amino acids which provide the ability for the GTP phosphates to cover the glycines of the GGDEF motif; this may be why the glycines within the GGDEF motifs are so highly conserved (Chan *et al.*, 2004). One GGDEF protein, ECA372, with a noncanonical SGDEF motif in *Pectobacterium atrosepticum* has been characterized as a diguanylate

cyclase. It is not yet known how this motif is able to bind to GTP; however, the ability of a noncanonical motif to bind GTP may indicate a flexibility of the binding site (Pérez-Mendoza *et al.*, 2011). For c-di-GMP to be formed, two GGDEF proteins must bind to GTP and then dimerize in an antiparallel fashion. This provides the right conformation for an intermolecular nucleophilic attack between one GTP and the α-phosphate group of the other GTP (Chan *et al.*, 2004, Whiteley & Lee, 2015). Once the pyrophosphate leaving group disperses, c-di-GMP leaves to elicit its responses.

The first EAL-domain phosphodiesterase to be crystallized was RocR of *Pseudomonas aeruginosa*. RocR has 8 alternating α helices and β strands along the backbone to form a triosephosphate isomerase (TIM)-barrel-like fold with the EAL (E175-A176-L177) domain located at the C-terminal part of the β barrel (Fig. 2) (Rao *et al.*, 2008, Tchigvintsev *et al.*, 2010). The c-di-GMP is bound within this barrel conformation and the secondary messenger is held in an extended conformation (Rao *et al.*, 2008, Tchigvintsev *et al.*, 2010, Bellini *et al.*, 2017). EAL proteins can either bind a Mg^{2+} or manganese (Mn^{2+}) to conduct phosphodiesterase activity, with Mg^{2+} thought to be more physiologically relevant (Schmidt *et al.*, 2005, Rao *et al.*, 2008). This Mg^{2+} ion coordinates an H_2O molecule to be linear with the electrophilic phosphorous of c-di-GMP. The H_2O molecule is deprotonated to produce a nucleophilic hydroxide ion. The deprotonated water engages in a nucleophilic attack onto the electrophilic phosphorous of c-di-GMP. This produces 5'pGpG, a linear nucleotide (Rao *et al.*, 2008).

HD-GYP domains are the second class of PDEs that break down c-di-GMP. They hydrolyze c-di-GMP into 2 GMP molecules. HD-GYPs are part of a large HD domain

subfamily of phosphomonoesterases and phosphodiesterase which use mono- or binuclear catalytic metal center to catalyze their respective reactions (Bellini *et al.*, 2014).

These c-di-GMP associated proteins have a high degree of redundancy, especially in Gram-negative bacteria. *Vibrio cholerae*, for instance, has 62 c-di-GMP associated proteins: 31 GGDEF proteins, 12 EAL proteins, 9 HD-GYP proteins, and 10 tandem GGDEF-EAL proteins. Some of this redundancy can be explained with these proteins commonly being linked with sensory domains. These sensory domains modulate enzymatic function in accordance to environmental signals (Karatan & Watnick, 2009, Whiteley & Lee, 2015). There is also evidence for different enzymes having different rates of input and output to enable fine tuning of the phenotypes. This was established in diguanylate cyclases in *V. cholerae* as they can be characterized as either fast biofilm regulators or slow biofilm regulators based on how fast these enzymes affect *vps* gene transcription (Massie *et al.*, 2012).

Tandem GGDEF-EAL proteins

Tandem GGDEF-EAL proteins are seen in a multitude of bacterial species; a recent genomic analysis found that approximately one third all of proteins with GGDEF domains and approximately two thirds of those with EAL domains contained both GGDEF and EAL domains in tandem. In *V. cholerae*, approximately one sixth of the GGDEF/EAL proteins contain both domains. Despite the abundance of tandem GGDEF-EAL proteins, they are not well characterized; of the 4867 such proteins, only a handful have been characterized (Römling *et al.*, 2013). The majority of tandem GGDEF-EAL proteins have been characterized in other bacterial species besides *V. cholerae*. It is important to note that the tandem GGDEF-EAL proteins do not necessarily indicate both diguanylate cyclase and

phosphodiesterase activity. Only three proteins have been characterized as dual-functional diguanylate cyclases and phosphodiesterases in bacteria: DcpA of *Agrobacterium tumefaciens,* ScrC of *Vibrio parahaemolyticus,* and BphG1 of *Rhodobacter sphaeroides* (Tarutina *et al.*, 2006, Ferreira *et al.*, 2008, Feirer *et al.*, 2015). Many tandem GGDEF-EAL proteins have been characterized as active phosphodiesterases; in this case, the GGDEF domains of these proteins have either a regulatory role or are uncharacterized. For example, the *C. crescentus* GGDEF-EAL protein CC3396 is a phosphodiesterase with a noncanonical GGDEF domain (GEDEF) that can bind GTP. This binding of GTP activates the phosphodiesterase activity in CC3396; however, the physiological relevance of this activation is not currently understood. It is possible that a decrease in GTP concentration could be a stress response and indicates the need to turn off phosphodiesterase activity to avoid depleting the GTP pool. Another possibility is that a decrease in GTP would inhibit phosphodiesterase activity, increasing c-di-GMP concentration (Christen *et al.*, 2005). Several other tandem GGDEF-EAL proteins have been characterized to have phosphodiesterase activity; however, the role of the GGDEF domains within the proteins remain elusive. BifA of *Pseudomonas aeruginosa*, for example, is another GGDEF-EAL protein that has been characterized as a phosphodiesterase as it suppresses biofilm formation. Interestingly, the biofilm reactions require both active site motifs as alanine substitutions in either the GGDEF (AAAAF) or EAL (AAL) fail to complement the biofilm defect of the Δ*bifA* mutant. How the GGDEF domain is necessary has yet to be discovered (Kuchma *et al.*, 2007).

The biofilm-regulatory GGDEF-EAL protein MbaA

One of these tandem GGDEF-EAL proteins is the biofilm regulator MbaA of *V. cholerae*. MbaA is bound to the membrane by two transmembrane regions with a periplasmic region between them. The GGDEF and EAL domains are within the cytoplasmic region. MbaA has previously been characterized as a phosphodiesterase (Bomchil *et al.*, 2003, Cockerell *et al.*, 2014). However, the function of the GGDEF domain has yet to be elucidated. The GGDEF domain has the potential of being an active diguanylate cyclase but could also play a regulatory role within MbaA.

MbaA works within the NspS-MbaA signaling complex. NspS is a periplasmic protein that binds and responds to polyamines (Cockerell *et al.*, 2014, Sobe *et al.*, 2017). Polyamines are ubiquitous molecules made up of carbon chains with two or more amine groups (Karatan and Watnick 2009). NspS responds to three polyamines: norspermidine, spermidine, and spermine (McGinnis *et al.*, 2009, Cockerell *et al.*, 2014, Sobe *et al.*, 2017). Norspermidine is found in the aquatic environment and enhances *V. cholerae* biofilm formation (Karatan *et al.*, 2005). NspS is hypothesized to bind to norspermidine and then interact with the periplasmic region of MbaA and inhibit phosphodiesterase activity, allowing for the accumulation of c-di-GMP (Karatan *et al.*, 2005, Cockerell *et al.*, 2014). Spermine and spermidine are found in foodstuffs and produced by the human microbiome. It is hypothesized that when NspS senses spermine and spermidine *V. cholerae* reduces biofilm formation by inhibiting the interaction between NspS and MbaA (McGinnis *et al.*, 2009, Sobe *et al.*, 2017). If these hypotheses are true, these polyamines could be environmental signals that *V. cholerae* senses in order to recognize whether it is in a human host or in its

aquatic environment. When *V. cholerae* senses a change in environment, it would need to establish what state to be in: planktonic or biofilm (Karatan *et al.*, 2005, Sobe *et al.*, 2017).

This NspS-MbaA signaling pathway may be a putative novel signaling pathway found in different bacteria. When a comparative genome analysis was conducted within sequenced bacterial genomes, it was revealed that 13 Proteobacteria species have periplasmic binding protein-encoding genes potentially co-transcribed with genes encoding MbaA-like tandem GGDEF-EAL proteins (Table 1) (Cockerell *et al.*, 2014). This would indicate that characterizing the NspS-MbaA pathway may not only apply to *Vibrio cholerae* but yield insights into c-di-GMP signaling in many other bacteria.

Table 1. NspS-MbaA gene pairs in other bacteria (Cockerell *et al.* 2014)**.**

Organism	PBP	Predicted ligand	GGDEF/EAL	Proximity
Vibrio cholerae	*nspS*	Norspermidine/spermidine	*mbaA*	Overlap
Psychromonas ingrahamii	Ping_1238	Spermidine/putrescine	Ping_1239	8
Hahella chejuensis	HCH_06688	Spermidine/putrescine	HCH_06689	4
Shewanella sediminis	Ssed_2394	Spermidine/putrescine	Ssed_2393	Overlap
Pseudomonas stutzeri	PST_0371	Spermidine/putrescine	PST_0370	13
Sinorhizobium meliloti	SMc_00991	Putrescine	SMc_00991	3
Magnetospirillum magneticum	amb_1105	Phosphate/phosphonate	amb_1104	3
Nitratiruptor sp. SB155-21	NIS_1757	Nitrate/sulphonate/bicarbonate	NIS_1758	Overlap
Thiomicrospira crunogena	Tcr_1221	Phosphonate	Tcr_1222	Overlap
Vibrio parahaemolyticus	VPA_1753	Alkylphosphonate	VPA_1754	Overlap
Vibrio parahaemolyticus	VPA_1512 (ScrB)	S-signal	VPA_1511 (ScrC)	Overlap
Shewanella paeleanna	Spea_3650	ESBF-3	Spea_3649	13
Xanthomonas oryzae pv. *oryzae*	X000RF_2004	Phosphate	X000RF_2004	Overlap
Xanthomonas campestris pv. *campestris*	xcc-b100_1903	Anions	xcc-b100_1904	3

My research aims to better characterize the NspS/MbaA signaling pathway. The first step is to understand the role of the GGDEF domain in MbaA. This research is important to better understand how *V. cholerae* modifies its biofilm formation through the control of c-di-GMP. If the GGDEF domain is a functional diguanylate cyclase, then the characterization of MbaA will give insight into the function of other dual function diguanylate cyclases phosphodiesterase enzymes within the bacterial world. If the GGDEF domain has a

regulatory effect on the EAL domain, this may shed light in how phosphodiesterase activity is controlled in other tandem GGDEF-EAL proteins. Both of these results would give insight into exploring the other 9 tandem GGDEF-EAL proteins found in *V. cholerae* and the 13 other bacterial species with NspS-MbaA like protein pairs.

Cholera takes 100,000 lives worldwide every year; it is of great importance to understand the life cycles of pathogenic bacteria in hopes of finding a key to treat or prevent infections. *V. cholerae* biofilms are believed to play a part in the pathogenesis of cholera; therefore, understanding how biofilms are regulated could give us clues to interrupt cholera pathogenesis.

Materials and Methods

Bacterial strains and growth conditions

E.coli and *V. cholerae* strains, plasmids, and primers used in this study are listed in Table 2 and Table 3, respectively. Primers were synthesized by Sigma-Aldrich (St. Louis, MO). Sequencing was performed by Eurofins (Louisville, KY). All strains were grown in Luria Bertani broth (LB) medium (per liter: 10 g tryptone, 5 g yeast extract, 10 g NaCl) with appropriate antibiotics. Sucrose agar (333.3 mL of DI water, 2.5 g of yeast extract, 2.5 g of tryptone, 7.5 g of agar, 30% sucrose) plates were used when mentioned. *E. coli* strains were incubated at 37°C and *V. cholerae* was incubated at 27°C, except *V. cholerae* strains containing pEVS143 or pMbaA, which were incubated at 37°C unless otherwise specified. Ampicillin, streptomycin, and kanamycin were used at a concentration of 100 µg/mL unless otherwise specified.

Multiple Sequence Alignment of GGDEF-EAL proteins

ScrC and DcpA are hypothesized dual functional diguanylate cyclase phosphodiesterases and were chosen as comparisons to MbaA amino acid sequence alignment (Ryjenkov *et al.*, 2005, Ferreira *et al.*, 2008, Feirer *et al.*, 2015). Sequences were retrieved from the Integrated Microbial Genomes and Microbiomes database (https://img.jgi.doe.gov/): MbaA (Q9KU26), DcpA (NP_357114), and ScrC (AAK08640). These sequences were input into the Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de/) to identify the GGDEF-EAL regions of the proteins. The

regions of the proteins containing only the GGDEF and EAL domains were then aligned using the multiple sequence alignment tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

Table 2. Bacterial strains and Plasmids used in this study[a]

Strains or plasmids	Description	Source
E. coli strains		
DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Invitrogen
NEB express	*fhuA2* [*lon*] *ompT gal sulA11 R(mcr-73::minutesiTn10*--Tet^S)*2 [dcm] R(zgb-210::Tn10*--Tet^S) *endA1 Δ(mcrC-mrr)114::IS10*	New England Biolabs
DH5α λ pir	*sup*E44 Δ*lac*U169 *hsd*R17, *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1, λpir	(Hanahan, 1983)
SM10 λpir	*thi thr leu tonA lacY supE* recA::*RP4-2-Tc::MuλpirR6K*; Km^R	(Moorthy & Watnick, 2005)
AK730	DH5α with pWHIT1; Ap^R	This study
AK737	DH5α with pWHIT2; Ap^R	This study
AK747	DH5α with pWHIT3; Ap^R	This study
AK756	NEB express with pWHIT4; Ap^R	This study
AK805	DH5α with pWHIT5	This study
AK338	NEBexpress with pRC2	(Cockerell *et al.* 2014)
AK806	DH5α with pWHIT6	This study
AK817	DH5α with pWHIT7	This study
AK811	DH5α with pWHIT8	This study
AK833	Sm10λpir with pWHIT7	This study
AK842	Sm10λpir with pWHIT8	This study
NEB α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	(Lee *et al.*, 2015)
AK856	NEB 5α with pWHIT9	This study
AK868	NEB 5α with pWHIT10	This study
V. cholerae strains		
PW357	MO10*lac*Z::*vpsL*p→*lac*Z, Sm^R	(Haugo & Watnick, 2002)
AK839	PW357 Δ*mbaA-GGDEF*	This study

Strains or plasmids	Description	Source
V. cholerae strains		
AK845	PW357 Δ*mbaA-EAL*	This study
PW444	MO10*lacZ*::*vpsL*p→*lacZ* Δ*mbaA*, Sm^R	(Bomchil *et al.*, 2003)
PW514	MO10*lacZ*::*vpsLp*→*lacZ* Δ*nspS*, Sm^R	(Karatan *et al.*, 2005)
AK555	PW444 with pMbaA, Km^R	(Sobe *et al.*, 2017)
AK865	PW444 with pWHIT9, Km^R	This study
AK872	PW444 with pWHIT10, Km^R	This study
AK564	PW444 with pEVS143	
Plasmids		
pCR2.1-TOPO	Plasmid for TOPO cloning, Ap^R	Invitrogen
pMAL-c5x	*malE* encoded with in frame insertion of the C-terminutesus of domain; Ap^R	New England Biolabs
pWHIT1	pCR2.1-TOPO::Mba-GGDEF domain; Ap^R	This study
pWHIT2	pCR2.1-TOPO::Mba-EAL domain; Ap^R	This study
pWHIT3	pMAL-c5x::Mba-GGDEF domain; Ap^R	This study
pWHIT4	pMAL-c5x::Mba-EAL domain; Ap^R	This study
pWHIT5	pCR2.1-TOPO::MbaAΔEAL, Ap^R	This study
pWHIT6	pCR2.1-TOPO::MbaAΔGGDEF, Ap^R	This study
pWM91	*oriR6k*, *lac*Aα, *sacB*, homologous recombination plasmid, Ap^R	(Metcalf *et al.*, 1996)
pWHIT7	pWM91:: MbaAΔGGDEF, Ap^R	This study
pWHIT8	pWM91:: MbaAΔEAL, Ap^R	This study
pEVS143	Km^R	(Dunn *et al.*, 2006)
pMbaA	pEVS143::VC0703, Km^R	(Massie *et al.*, 2012)
pWHIT9	pMbaAΔEAL, Km^R	This study
pWHIT10	pMbaAΔGGDEF, Km^R	This study

[a]Anibiotic resistance phenotypes: streptomycin, Sm^R; ampicillin, Ap^R; kanamycin, Km^R

Table 3. Primers used in this study*

Primers	Description	Sequence (5'→3')	Source
M13 forward	Used for sequencing insertion into pCR2.1-TOPO and pMAL-c5x	GTAAAACGACGG CCAG	(Messing, 1983)

Primers	Description	Sequence (5'→3')	Source
M13 Reverse	Used for sequencing insertion into pCR2.1-TOPO and pMAL-c5x	CAGGAAACAGCTATGAC	(Messing, 1983)
PA323	Forward primer for the cloning of MbaA-GGDEF in to pMAL-c5x with a NdeI site	CATATGGCCATGTACAGCGAACTTC	This study
PA324	Reverse primer for the cloning of MbaA-GGDEF in to pMAL-c5x with a BamHI site	GGATCCCTAATAGTAAGCGACTTGGTT	This study
PA325	Forward primer for the cloning of MbaA-EAL in to pMAL-c5x	CATATGCAGCGGCGCAATAATATTG	This study
PA326	Reverse primer for the cloning of MbaA-EAL in to pMAL-c5x	GGATCCCTAACGGCATTCACTTTG	This study
PA335	Reverse primer for upstream of GGDEF deletion	CGCAAGTGCTTTCGTGCG	This study
PA336	Forward primer for upstream of GGDEF deletion	CTCAGTTATAACCGCCAAGACGTG	This study
PA337	Forward primer for downstream of GGDEF deletion with SOE tag	CGCACGAAAGCACTTGCGGTCGCTTACTATTCGCAG	This study
PA338	Reverse primer for downstream of GGDEF deletion	CGGCAAACGTTTCGGTGA	This study
PA339	Reverse primer for upstream of EAL deletion	ATTGCGCCGCTGTACTATTTG	This study
PA340	Forward primer for upstream of EAL deletion	GATACTCAGCATATCTGGGTCATG	This study
PA341	Forward primer for downstream of EAL deletion with SOE tag	CAAATAGTACAGCGGCGCAATTAGAAGTGCCTTTAGCACCG	This study
PA342	Reverse primer for downstream of EAL deletion	AGAGCGAGAATTAAGGCCTGA	This study

Primers	Description	Sequence (5'→3')	Source
PA364	Reverse primer for amplification of the end region of MbaA	CTAACGGCATTCACTTTGGCTGGG	This study
PA370	Forward primer for amplification of GGDEF deletion for use in Q5 mutagenesis kit	TCGCAGGCGTTAGATCAAATA	This study
PA371	Reverse primer for amplification of GGDEF deletion for use in Q5 mutagenesis kit	TTCGCTGTACATGGCGTA	This study
PA372	Forward primer for amplification of EAL deletion for use in Q5 mutagenesis kit	ATCTCCCCATGGCTTCATG	This study
PA373	Reverse primer for amplification of EAL deletion for use in Q5 mutagenesis kit	GCGCCGCTGTACTATTTG	This study

*Restriction enzyme sites are underlined

Cloning regions encoding the individual domains of MbaA into pMAL-c5x

Primers were designed that flanked the regions of the gene encoding individual domains of MbaA. Domains were identified using SMART. PA323 and PA324 had a melting temperature of 66 °C and 67.5 °C, respectively. They were designed to amplify a 570 bp fragment that encoded amino acids 324 through 502 from the end of the transmembrane region to the beginning of the EAL domain, which would capture the GGDEF domain in its entirety. PA325 and PA326 had a melting temperature of 64.6 °C and 67.6 °C, respectively. They were designed to amplify an 843 base pair fragment that encoded amino acids 515-756, from the end of the GGDEF domain to the end of MbaA, which encompasses the EAL domain. Restriction enzyme cut sites of BamHI and NdeI were introduced into the 5' ends of the primers for later use. PCR was run using Q5 polymerase (New England Biolabs, Inc,

Ipswich, MA). Cycling conditions were as follows: initial denaturation at 98 °C for 30 seconds; 35 cycles of denaturation at 98 °C for 10 seconds, annealing at 59 °C for 30 seconds, and extension at 72 °C for 30 seconds; final extension of primers was set at 72 °C for two minutes. Adenines were added to 3' blunt ends of the PCR products by following the manufacturer's instructions: 10 µL of PCR product was mixed with 33.5 µL water, 5 µL OneTaq polymerase buffer, 1 µL dATP (10 mM), and 5 µL OneTaq polymerase and incubated at 72 °C for 10 minutes. This product was then cloned into the pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA), which contains thymine overhangs on the linearized plasmid. This produced pWHIT1 and pWHIT2 for the GGDEF domain and EAL domain, respectively (Table 2). These plasmids were then electroporated into the *Escherichia coli* DH5α using a BIO-RAD MicroPulser (Hercules, CA) at 1.8 kV. The transformed DH5α was recovered with SOC media at 37 °C at 200 rpm for one hour. The cells were then plated onto LB agar with 100 µg/L amp and 40 µg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X Gal). Blue-white screening was conducted to find colonies containing plasmids with the cloned inserts. White colonies were chosen to use in a colony PCR. Single cells were suspended in 100 µL of water and then lysed at 95 °C for 5 minutes. Two µL of this lysate was then used as a template in a PCR reaction consisting of 16.4 µL water, 5 µL 5X OneTaq Standard Reaction Buffer, 0.5 µL dNTPS, 0.5 µL of each 10mM forward and reverse primer, and 0.125 µL OneTaq Hot Start DNA polymerase (New England Biolabs, Inc, Ipswich, MA). Plasmids were extracted from three colonies containing the confirmed insert using a Wizard Plus SV Miniprep DNA Purification Kit (Promega, Madison, WI). The

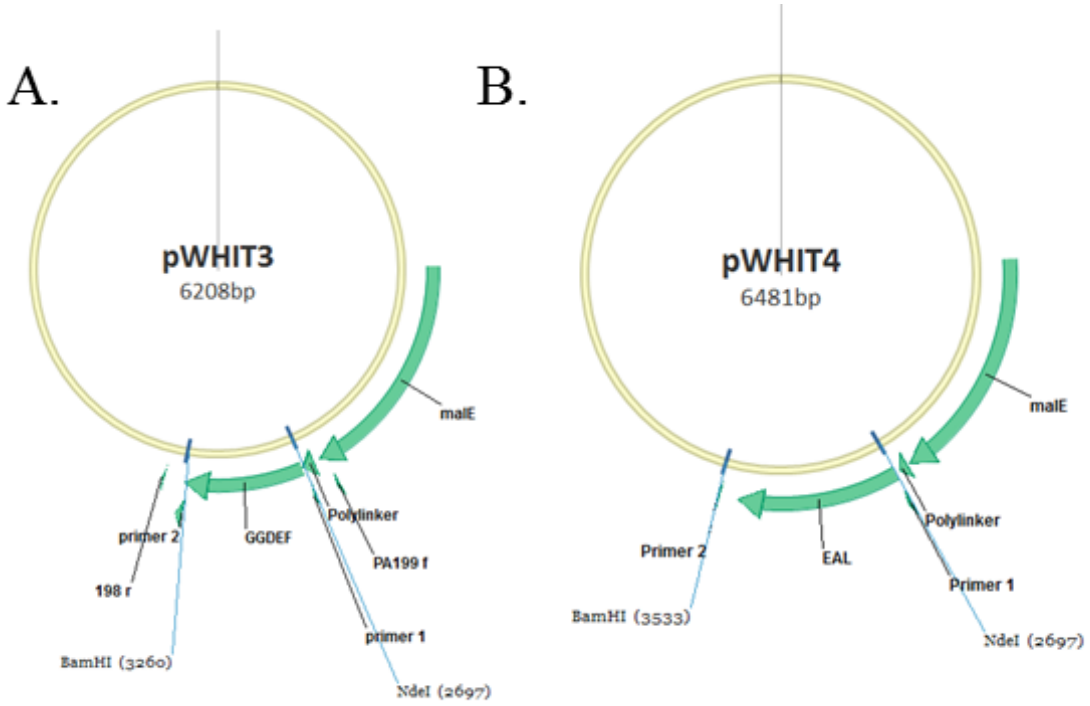

Fig 2. Plasmid map of (a) pWHIT3 and (b) pWHIT4. (a) pWHIT3 and (b) pWHIT4 was constructed by digesting pMal-c5x and pWHIT2 and pWHIT1, respectively, with BamHI and NdeI and then ligating with Electroligase. Domains were designed to be in frame of the *malE* gene which encodes a maltose binding protein.

plasmids were further confirmed by sequencing using forward and reverse M13 primers (Table 2).

The plasmids pWHIT1 and pWHIT2 were then digested using BamHI and NdeI to release the insert. The expression plasmid pMAL-c5x (New England Biolabs, Inc), which encodes a maltose binding protein from *E. coli*, was also digested with the same enzymes and the insert was ligated into pMAL-c5x to form pWHIT3 and pWHIT4, respectively, using Electroligase (New England Biolabs, Inc.) (Fig. 2; Table 2). The product was designed so that the produced protein would be in frame with the encoded modified maltose binding protein (*malE*). This produces a fusion between the maltose binding protein and the protein domain of interest. The maltose binding protein has been modified to bind to amylose as a

way to purify the protein of interest. The plasmids pWHIT3 and pWHIT4 were electroporated into DH5α and recovered as previously described (Table 1). To confirm the insertion was present colony PCR was performed and plasmid was extracted as described above. The plasmid was then chemically transformed into NEB express (New England Biolabs, Inc) by incubating on ice for 15 minutes, then a 42 °C water bath for 65 seconds and then placed on ice for three minutes. Cells were recovered in LB media for one hour at 37 °C at 200 rpm. NEB express transformed with pWHIT3, which encoded the GGDEF domain, did not survive. Diguanylate cyclases have been documented to be lethal in some *E. coli* strains. Other studies have previously cloned diguanylate cyclases using the pMAL systems in DH5α; therefore, pWHIT3 was retained in DH5α while pWHIT4 was successfully transformed into NEB express (10).

Expression and purification of MBP-GGDEF and MBP-EAL

All proteins purified in this research had a pMAL backbone and were purified in the same way unless otherwise specified. The fusion proteins were expressed in their respective *E. coli* strain at 37°C for 17 hours 10 mL LB with 100 µg/mL ampicillin and 0.2% glucose; glucose inhibits amylase production which would degrade the amylose resin used during purification. MBP-Cterm is a fusion protein of maltose binding protein and the C terminus portion of MbaA containing both the GGDEF and EAL domains; this protein was used previously to characterize MbaA phosphodiesterase activity (Cockerell *et al.*, 2014). Slr1143 is a maltose binding protein fusion with the *Synechocystis sp. PCC6803* protein Slr1143 which has been characterized as an active diguanyalte cyclase (Ryjenkov *et al.*, 2005). The 10mL overnight cultures were then added to 1L cultures of LB with 100 µg/mL ampicillin and 0.2% glucose and *E. coli* cells were grown to an OD_{595} of 0.4 to 0.6. Expression was

induced with the addition of IPTG to the final concentration of 0.3 mM. Cells were grown overnight at respective temperatures with shaking at 200 rpm and then centrifuged at 4,000 x g for 20 minutes at 4 °C. Cells expressing MBP-GGDEF, MBP-EAL, and Slr1143 were incubated at 30°C while cells expressing MBP-Cterm was incubated at 18°C to decrease inclusion body formation. The cells were resuspended in 25 mL column buffer (200 mM NaCl, 0.5 mM EDTA, 5 mM MgCl2, 20 mM Tris-HCl (pH 7.6)) and then frozen. The cells were thawed on ice and then sonicated for 15-second intervals for a total of 2 minutes. Lysates were then centrifuged at 20,000 x g for 20 minutes at 4°C to pellet cells. The cleared lysate was then transferred to a 50 mL conical tube and combined with the prepared amylose resin. One mL of amylose resin was prepared by resuspending the resin with column buffer and then centrifuged at 700 x g for 1 minutes, this process was repeated three times producing 500 µL of packed amylose resin. The amylose resin lysate mix was then incubated overnight at 4 °C with rotation. The protein was then washed with 50 mL of column buffer and then eluted four times with 1 mL of elution buffer (column buffer with 20mM maltose) using a gravity column. Ten µL of each fraction were run on a 10 % SDS-PAGE gel at 250 V for 45 minutes and stained with 0.01% coomasie brilliant blue.

Phosphodiesterase and diguanylate cyclase assays

Purified protein fractions were combined and were dialyzed into their respective reaction buffer using Slide-A-Lyzer dialysis cassettes (Pierce) two times for 4 to 16 hours at 4 °C. MBP-EAL was dialyzed into a phosphodiesterase buffer (50mM Tris, adjusted to pH 8.5) and the MBP-GGDEF was dialyzed into diguanylate cyclase buffer (50 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 0.5 mM EDTA, and 50 mM NaCl). The protein was extracted from cassettes and concentration calculated using BSA standard curve; this BSA standard

curve was calculated separately with each buffer to provide accurate results. $MnCl_2$ was added to the phosphodiesterase reaction after dialysis at a final concentration of 2mM. The reactions contained 2.5 µM of MBP-GGDEF and MBP-EAL and 2 µM of MBP-Cterm. The substrate was added at 100 µM to start the reaction. Phosphodiesterase reactions were incubated at 37 °C for 2 hours. Diguanylate cyclase were incubated at 37 °C for 5 hours. A time course phosphodiesterase assay was set up as previously described with MBP-EAL and MBP-Cterm for 0, 30, 60, and 120 minutes. Reactions were stopped by the addition of $CaCl_2$ at the final concentration of 10 mM. $CaCl_2$ is a competitive inhibitor that prevents the binding of $MnCl_2$, which stops phosphodiesterase reactions.

HPLC analysis

The enzymatic reactions were boiled for 5 minutes and then centrifuged at 14,000 x g for 7 minutes using a Nanosep 10 kDa Omega filter (Pall Life Sciences, Port Washington, NY) to remove proteins. Forty µL of the reaction products were separated using a SUPELCOSIL™ LC-18 column (Sigma-Aldrich, St. Louis, MO) with a Waters 1525 Binary HPLC Pump and analyzed using a Waters 2487 Dual λ Absorbance detector (Milford, MA). Peaks were identified by comparison to purified GTP, c-di-GMP, and pGpG standards (Biolog Life Science Institute, Brememn, Germany; Jena Bioscience, Jena, Germany). The areas under the curves representing c-di-GMP and 5' pGpG peaks were used to calculate product formation and was analyzed using Microsoft Excel (Redmond, WA).

Deletions of the individual domains of MbaA on V. cholerae chromosome

In order to characterize how these domains, regulate biofilm formation I decided to delete the individual domains from the chromosome. The strains carrying these domain deletions would then be used in a biofilm assay to see if these domains play a part in biofilm

formation. Primers with complementary regions were designed to flank regions of the *mbaA* gene encoding the GGDEF and EAL domains (Table 3). PA335 and PA336 had a melting temperature of 58 °C and 57.6 °C, respectively, and were designed to amplify 477 bp between the amino acids 180 and 333, spanning a portion of the transmembrane region before the GGDEF domain (Fig. 3). PA337 and PA338 had a melting temperature of 67.9 °C and 56.4 °C, respectively, and were designed to amplify 460 bp between the amino acids 498 and 645, spanning a fragment of the EAL domain (Fig. 3). These two fragments were designed to flank the GGDEF domain (Fig. 3). PA339 and PA340 had a melting temperature of 56.1 °C and 54.8 °C, respectively, and were designed to amplify 463 bp region between the amino acids 360 and 512, spanning a portion upstream of the EAL domain (Fig. 4). PA341 and PA342 had a melting temperature of 66.6 °C and 55.8 °C and were designed to amplify 462 bp which is between 15 bp upstream of the stop codon of MbaA and 419 bp down stream of MbaA stop codon, spanning the last portion of MbaA and an area downstream (Fig. 4). These two fragments were designed to flank the EAL domain (Fig. 4). The resulting PCR products were joined together with the help of the complementary regions by overlap extension PCR. The resulting construct was TOPO cloned into pCR2.1-TOPO as described above (Table 2). Correct construction was confirmed by colony PCR and sequencing. The plasmid was then extracted and purified as previously described. The plasmid and the suicide vector pWM91 were digested using the restriction enzyme XhoI and SpeI. Enzymes were heat killed at 80°C for 20 minutes. Both digested insert and plasmid were gel purified using the Monarch Gel Purification Kit (New England Biolabs, Inc). For the EAL deletion, pWM91 was also dephosphorylated using Antarctic phosphatase (New England Biolabs, Inc) to prevent religation of the vector and facilitate ligation of the ΔEAL

insert. The digested and gel purified pWM91 was mixed with 2 µL of 10x Antarctic phosphatase buffer and 5 µL Antarctic phosphatase and incubated at 37°C for 30 minutes. Enzyme was heat killed at 80°C for 2 minutes. The plasmid was ligated and then electroporated into an *E. coli* DH5α λ pir. To verify insertion, colony PCR was performed as previously described. The plasmid was then isolated as previously described and electroporated into Sm10 λ pir. This strain was used for conjugation with *V. cholerae* to generate in-frame chromosomal deletions through a SacB counter-selectable mutagenesis procedure (Metcalf *et al.*, 1996).

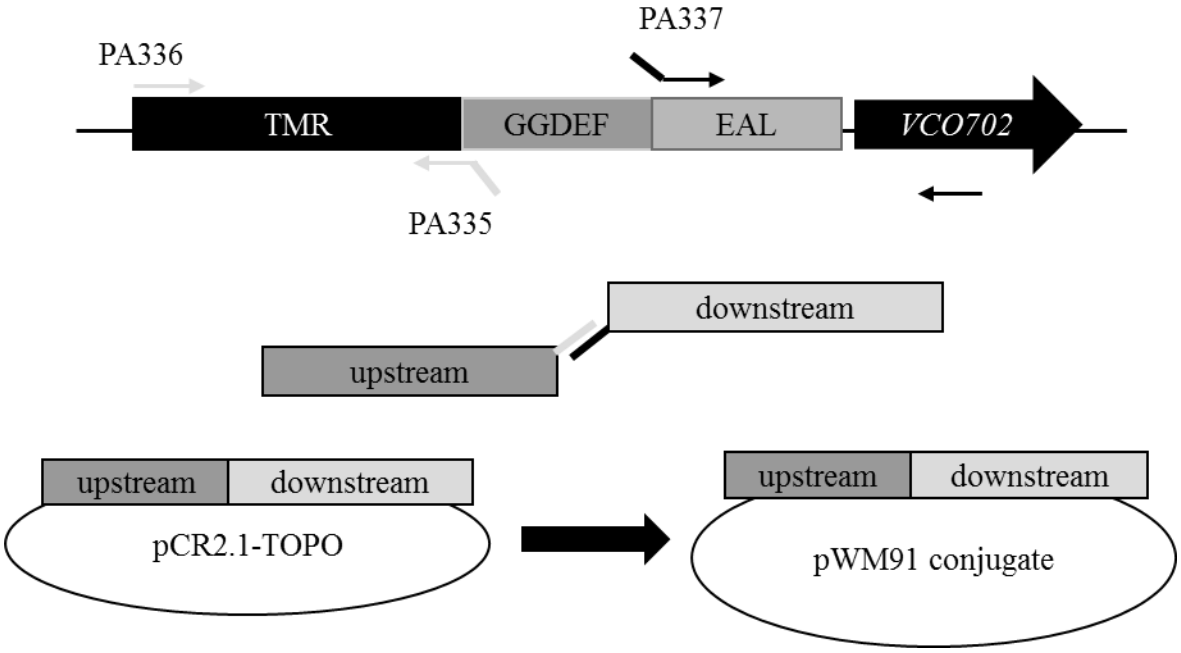

Fig 3. Construction of conjugation plasmid for the deletion of the GGDEF domain within MbaA. Primers were designed to amplify a ~400 bp upstream and downstream of the GGDEF domain. The forward primer of the downstream region was designed to be complementary to the reverse primer of the upstream region. Using overlapping PCR, the downstream and upstream were combined. This was then TOPO cloned into pCR2.1-TOPO to produce pWHIT6. It was then ligated into pWM91 using XhoI and SpeI to produce pWHIT7.

V. cholerae strain PW357 was used as a wild-type strain (Table 1). The respective *E. coli* SM10λpir strain was mated with PW357 (Table 2). Briefly, *V. cholerae* and *E. coli* were

grown on LB-Agar plates with antibiotics and half of the growth was then streaked onto selective media (LB media with 100µg/ml streptomycin and 50µg/ml ampicillin), resistant colonies were then purified by streaking again on the same selective media. The colonies were then streaked onto LB media with no antibiotics to facilitate an intrachromosomal recombination. Next, single colonies were streaked on sucrose agar plates. Sucrose resistant colonies were then patched for selection on two plates: one plate of streptomycin and another of streptomycin and 50 µg/mL ampicillin. Colonies sensitive to ampicillin were selected and analyzed using colony PCR to identify colonies with domain deletions.

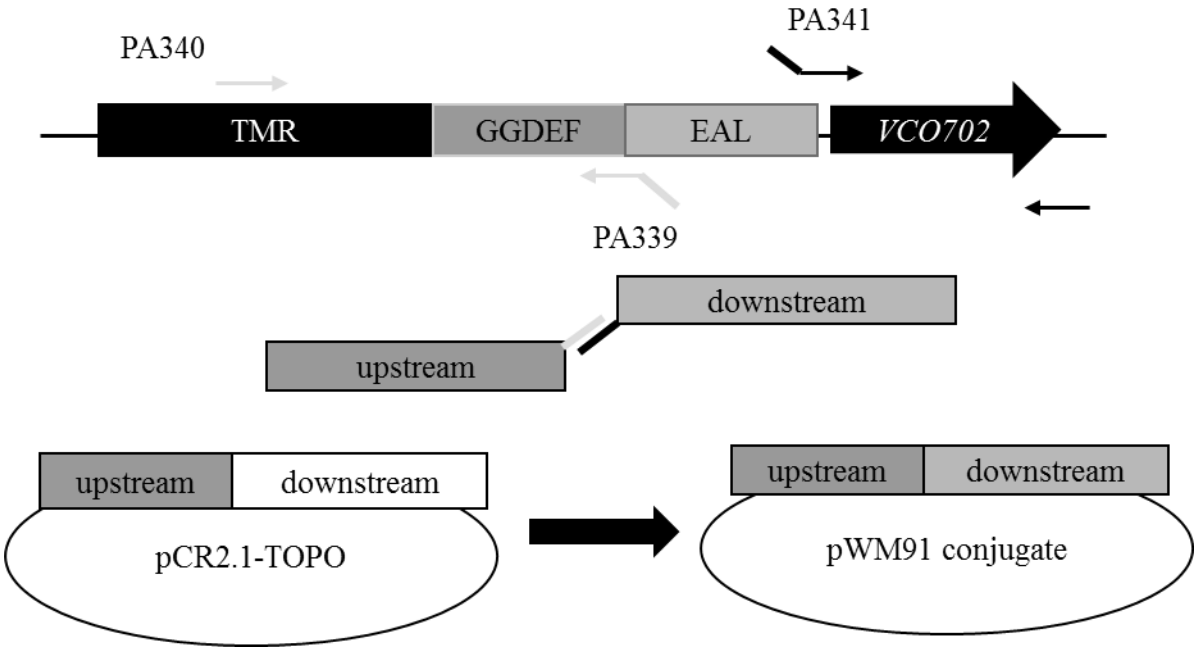

Fig 4. Construction of conjugation plasmid for the deletion of the EAL domain within MbaA. Primers were designed to amplify a ~400 bp upstream and downstream of the EAL domain. The forward primer of the downstream region was designed to be complementary to the reverse primer of the upstream region. Using overlapping PCR, the downstream and upstream were combined. This was then TOPO cloned into pCR2.1-TOPO to produce pWHIT5. It was then ligated into pWM91 using XhoI and SpeI to produce pWHIT8.

Deletion of regions encoding individual domains in pMbaA

Because stability issues were a concern after chromosomal deletions and because native levels of MbaA are too low to review in a western blot, domain deletions were performed on a plasmid. The Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Inc.) was used to create domain deletions of the plasmid pMbaA (Table 3). Primers were designed using NEBasechanger (New England Biolabs, Inc.) to amplify away from the domain of interest; this would lead to amplifying of the plasmid except for the domain of interest leading to a PCR product that includes the plasmid backbone and *mbaA* excluding a domain (Table 3; Fig. 5). Amplification using primers PA370 and PA371, which had a melting temperature of 64 °C resulted in deletion of amino acids 327 to 507 of MbaA. Amplification using primers PA372 and PA373, which had a melting temperature of 65 °C and 64 °C respectively, resulted in deletion of amino acids 518 to 760 of MbaA. Amplification was done using 12.5 µL 2x Q5 Master Mix (final concentration 1x), primers at final concentrations of 0.5 µM, and a pMbaA amount between 14 ng to 27.5 ng. PCR parameters were as follows: initial denaturation 98°C for 30 second, 30 cycles of 98°C 10 second, 59°C 30 second, 72°C 4 minutes, and a final extension of 72°C for 2 minutes. This PCR product was then used in a Kinase, Ligase, and DpnI (KLD) enzymes mix reaction (New England Biolabs, Inc.; Fig. 5). This mix is optimized to carry out multiple reactions in one tube. The kinase adds a phosphate group to the 5’ end of the amplicons, ligase ligates the completed PCR produce to circularize it and DpnI degrades the methylated template DNA (Fig. 4). The reaction contained 1 µL of PCR product, 5 µL 2x KLD reaction buffer (final concentration 1x), and 1µL 10x KLD enzyme mix (final concentration 1x). For the EAL deletion, the

reaction was set to 5 minutes at 21°C. The GGDEF deletion was set to 5 minutes at 21°C and then 30 minutes at 37°C to increase DpnI activity as initial attempts to delete the domain using the first set of conditions were unsuccessful. Five microliters of the reaction were then added to chemically competent NEB 5α cells (Table 1) for 30 minutes on ice. Cells were then heat shocked for 30 seconds and then placed on ice for 3 minutes. Cells were recovered in SOC media at 37°C for one hour. They were then plated on LB agar plates with kanamycin and incubated overnight. Colony PCR was used to confirm deletions using primers PA336 and PA364 (Table 2). These primers were also used to sequence the plasmids.

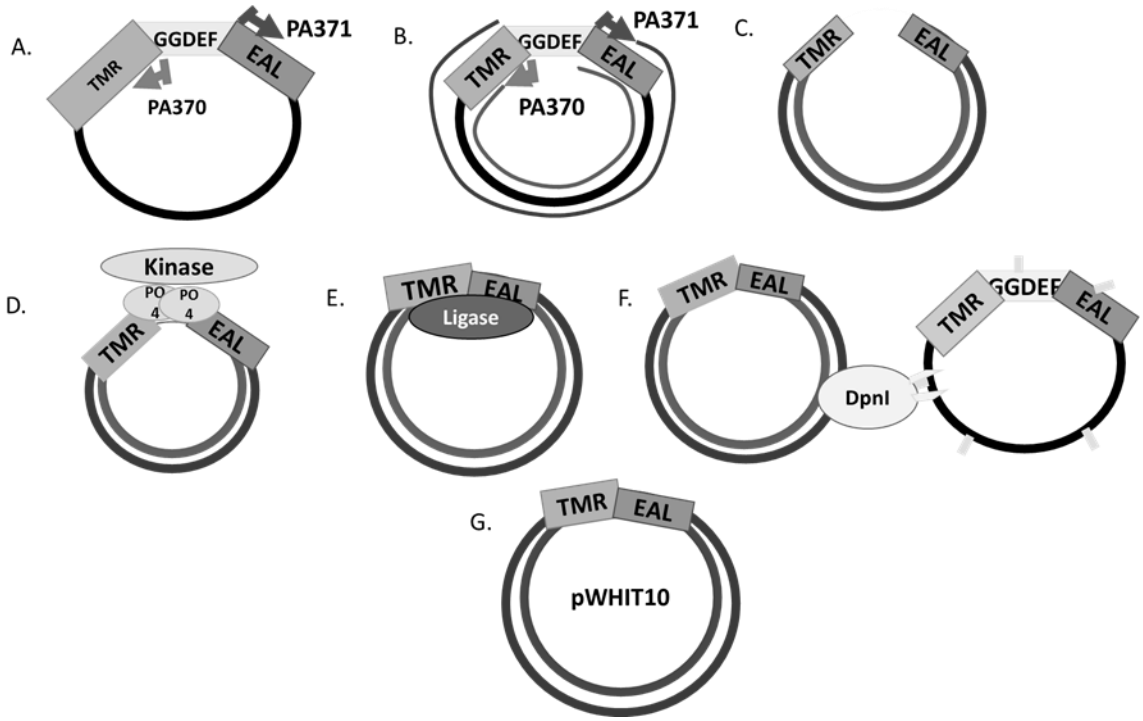

Fig 5. Depiction of Q5 mutagenesis procedure**.** (a) Primers were designed to amplify away from the GGDEF domain. (b) Amplification produced amplicons missing the GGDEF domain. (c) Hypothesized amplicon. (d) The amplicon was then used in a KLD reaction where first a kinase phosphorylates the 5' and 3' ends of the amplicon. (E) A ligase joins the two ends to produce a plasmid. (f) DpnI digests the methylated template plasmid (methyl groups shown in light gray). (g) Final product is pWHIT10 to be transformed into NEB 5α.

Biofilm assays for in vivo analysis of domain functions

A single colony was added to a 2 mL culture of LB with streptomycin and then allowed to grow for 17 hours (Table 2). A day culture was then set up with a 1:50 dilution of the original culture used to inoculate another 2 mL culture of LB with streptomycin and allowed to grow to an OD_{655} of 0.3-0.4. This was then normalized to an OD_{655} of 0.04 in a 300 µL culture inside a borosilicate tube. The bacteria were incubated for 16 hours at 27 °C. After incubation, the planktonic cells were removed and 150 µL was measured at OD_{595} in a 96-well microplate. The biofilm was then washed once with 300 µL of 1X phosphate-buffered saline (PBS). The biofilm was then homogenized by vortexing with 1.00 mm glass beads (BioSpec, Bartlesville, OK) in 300 µL of 1X PBS; 150 µL of this homogenized mix was added to the microplate and measured using a BIO-RAD iMark MicroPlate Reader (Hercules, CA) at a wavelength of 595 nm. Biofilm assays were performed with three technical replicates each time and repeated three times for accuracy. Student's t-tests were used to detect significant differences between various strains and treatments.

Results

Sequence analysis of the domains in MbaA

V. cholerae strictly regulates biofilm formation in response to environmental stimuli (Karatan & Watnick, 2009, Teschler *et al.*, 2015). The NspS-MbaA signaling system is hypothesized to be a regulation system that controls biofilm formation in response to specific polyamines (Karatan *et al.*, 2005, McGinnis *et al.*, 2009, Sobe *et al.*, 2017). The polyamines spermine and spermidine are sensed by NspS; this interaction is hypothesized to enhance the phosphodiesterase activity of MbaA which leads to a subsequent decrease in biofilm formation (McGinnis *et al.*, 2009, Sobe *et al.*, 2017). However, when NspS senses norspermidine, it is hypothesized to inhibit MbaA phosphodiesterase activity, allowing an accumulation of c-di-GMP, which enhances biofilm formation (Karatan *et al.*, 2005). MbaA harbors both GGDEF and EAL domains. Two other proteins, ScrC and DcpA, also have tandem GGDEF-EAL domains and have both diguanylate cyclase and phosphodiesterase activities (Ferreira *et al.*, 2008, Feirer *et al.*, 2015). While BphG1 is a dual functional diguanylate cyclase phosphodiesterase, it can only have diguanylate cyclase activity when the EAL domain has been removed (Tarutina *et al.*, 2006). Because BphG1 cannot function has a diguanylate cyclase in its complete form it was excluded from the alignment. This led me to hypothesize that MbaA is a dual-function enzyme with both diguanylate cyclase and phosphodiesterase activities. This would suggest that when NspS senses norspermidine it not

only inhibits phosphodiesterase activity, but possibly activates diguanylate cyclase activity, leading to increased c-di-GMP concentration in the cell.

To exhibit both activities, MbaA must have the catalytic amino acids necessary for both diguanylate cyclase and phosphodiesterase activity and have the flexibility to accommodate both activities. To determine whether MbaA had the conserved catalytic residues required for both activities, a multiple sequence alignment was done to compare the protein sequences of MbaA, DcpA, and ScrC (Fig. 6). I only analyzed the GGDEF-EAL portions of the proteins to ensure the best alignment (Fig. 6).

I wanted to first confirm that MbaA had all the catalytic amino acids necessary for phosphodiesterase activity. I located the catalytic amino acids first characterized in RocR of *P. aeruginosa*: Q161, E175, A176, L177, R179, E265, N233, D295, D296, and D318 (Rao *et al.*, 2008). I was able to identify all these catalytic amino acids within DcpA, ScrC, and MbaA. MbaA has the amino acids Q539 (RocR Q161), R558 (RocR R179), D673 (RocR D296), and D695 (RocR D318), which are required for interactions with the c-di-GMP molecule (Fig. 6) (Rao *et al.*, 2008). The amino acids that coordinate either the Mg^{2+} and Mn^{2+} ions were also identified in MbaA: E554 (RocR E175), N612 (RocR N233), D672 (RocR D295), E644 (RocR E265) (Fig. 6). The H_2O molecule would be coordinated by MbaA E727 (RocR E352) (Fig. 6). Without coordinating the metal ion and H_2O, there would be no nucleophilic attack on a phosphorous of c-di-GMP and it would not produce 5' pGpG. MbaA also had the amino acids V555 (RocR A176) and L556 (RocR L177), which stabilizes these proteins (Fig. 6) (Rao *et al.*, 2008). The valine to alanine substitution seen in Mba position V555 is a conserved substitution and seen in other phosphodiesterases (Kazmierczak *et al.*, 2006). Conservation of the amino acid residues required for

phosphodiesterase activity confirms that MbaA was correctly characterized as a phosphodiesterase.

I then analyzed if the GGDEF domain of MbaA had the potential to be a diguanylate cyclase. I used the catalytic amino acids first described in the characterized *C. crescentus* diguanylate cyclase PleD: R300, D327, N335, D344, G368, G369, E370, E371, F372, K442, R446 (Chan *et al.*, 2004). I then identified these in DcpA, ScrC, and MbaA. MbaA had the two amino acids necessary to bind GTP: N384 (PleD N335) and D393 (PleD D344) (Fig. 6). The magnesium ion would be coordinated by D376 (PleD D327), D426 (PleD E370), and E427 (PleD E371) (Fig. 6). The substitution of the aspartate (PleD E370) to glutamate (MbaA E427) within the GGDEF motif is a well-known substitution in diguanylate cyclases and should not affect activity (Whiteley & Lee, 2015). The intermediate formed in this reaction is stabilized by K496 (PleD K442) and K501 (PleD R446). The substitution of an arginine (PleD R446) to a lysine (MbaA K501) is a conserved substitution as both amino acids are basic; however, DcpA has a substitution of a proline (P) which is not a conserved substitution, this may indicate that this amino acid is not completely necessary for diguanylate cyclase activity (Fig. 6). The conserved sequence of GGDEF in MbaA is SGDEF (S424-G425-D426-E427-F428) (Fig. 6). The GTP in PleD is oriented to cover the conserved glycines; however, in MbaA the first glycine is substituted with a serine (Fig. 6). This was thought to produce an inactive GGDEF domain, but the *P. atrosepticum* ECA372 diguanylate cyclase also has a SGDEF domain which means a glycine to serine substitution within the GTP binding site can be accommodated by the enzyme (Pérez-Mendoza *et al.*, 2011). Therefore, this substitution does not rule out diguanylate cyclase activity. In the alignment,

MbaA is the only protein to have a seven amino acid insertion seven amino acids upstream of the GGDEF site. How this insertion affects MbaA activity is not known.

Fig 6. Multiple sequence alignment of MbaA, DcpA, and ScrC**.** The C-terminus sequences containing the GGDEF-EAL domains of MbaA, DcpA, and ScrC were aligned using Clustal Omega. The catalytic amino acids for diguanylate cyclase acitivity and phosphodiesterase activity were highlighted in grey. The amino acids between these domains are underlined in black. The seven amino acid insertion in MbaA is outlined in black.

There are several amino acids that link the GGDEF domain and EAL domain together. This linker is thought to be long enough to provide the protein with flexibility to accommodate both activities. I looked at the amino acids between these two domains to see if MbaA had similar number of amino acids as DcpA and ScrC. ScrC has 10 amino acids between its GGDEF and EAL domains and DcpA and MbaA have 9 (Fig. 6). Since MbaA has a similar number of amino acids between GGDEF and EAL to the other dual functional enzymes it is plausible that MbaA has enough structural flexibility to accommodate dual activities.

The GGDEF domain of MbaA does not have diguanylate cyclase activity in vitro

Because the multiple sequence alignment did not allow us to reject the hypothesis that the GGDEF domain could be a diguanylate cyclase, I wanted to test the GGDEF domain of MbaA for *in vitro* diguanylate cyclase activity. I designed primers to amplify the GGDEF domain of MbaA from chromosomal DNA (PW249) (Waldor & Mekalanos, 1994). This amplification produced a PCR product of approximately 0.5 kbp with some nonspecific amplification products of 0.4 kbp and 0.1 kbp (Fig. 7a). This product was gel purified and then cloned into pCR2.1 to produce pWHIT1. pWHIT1 was then sent out for sequencing to confirm correct sequence. The insert was then cloned into pMAL-c5x to produce pWHIT3. This plasmid was transformed into *E. coli* DH5α. Successful transformation was confirmed by colony PCR and the identification of an amplicon around 0.5 kbp (Fig. 7b). In colonies 1-4 a roughly 0.5 kpb insert can be seen confirming that the plasmid was successfully transformed into DH5α (Fig. 7b). For maximal expression of MBP-GGDEF, pWHIT3 was supposed to be transformed into NEB express *E. coli*; however, no transformants were obtained. It has been reported that diguanylate cyclases can be lethal to some strains of *E.*

coli, specifically BL21 derivatives, which is the type of strain NEB express was derived from. I, therefore, chose to express MBP-GGDEF in DH5α.

This protein was produced in 1 L cultures and purified using an affinity column with amylose-resin (Fig. 8). The maltose binding protein encoded in pMAL-c5x had been modified to bind to amylose. The fusion protein binds to amylose resin which allows contaminating proteins to be washed away as seen in the unbound fraction (UF) and wash (W) (Fig. 8). MBP-GGDEF was then eluted using 20 mM maltose which is seen in lanes 1-4 (Fig. 8).

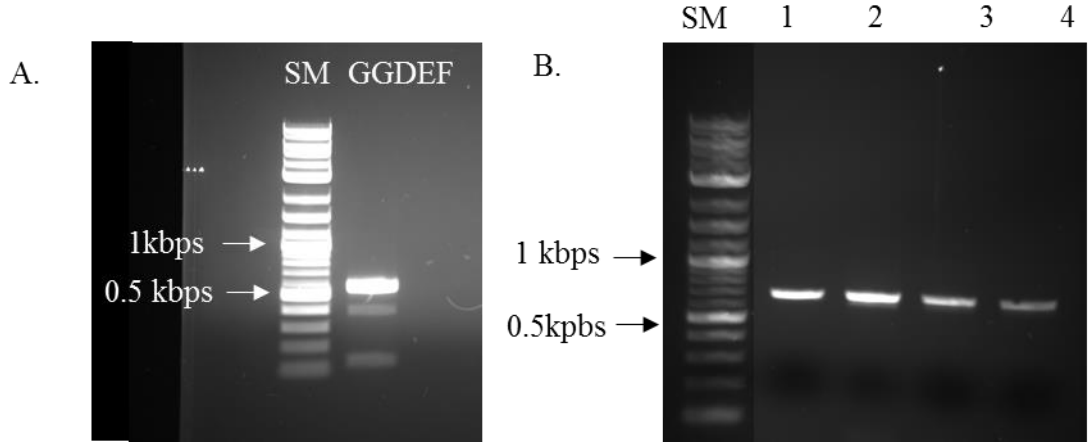

Fig 7. Construction of pWHIT3. (a) Amplification of the GGDEF domain from chromosomal DNA. The GGDEF amplicon was approximately 0.5 kbps based on size marker (SM). (b) Colony PCR of pWHIT3 within DH5α.

Maltose outcompetes amylose for binding to the maltose binding protein which allows the MBP fusion protein to be eluted (Fig. 8). The MBP-GGDEF protein is approximately 60 kDa. After the protein was confirmed to be eluted by SDS gel, fractions 1-4 were pooled, and then dialyzed into diguanylate cyclase reaction buffer (Fig. 8).

A diguanylate cyclase assay was set up as described in the methods and the products were then separated using HPLC. The peaks of the chromatograms were compared to GDP,

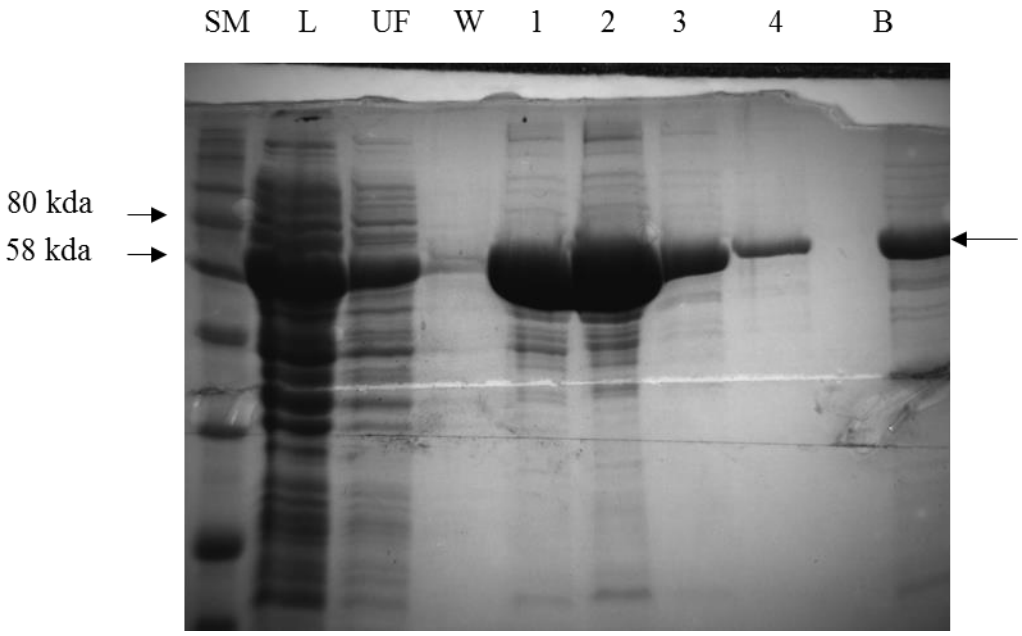

Fig 8. Purification of MBP-GGDEF**.** Lysate (L) from IPTG induced MBP-GGDEF in DH5α was collected before binding with maltose resin. The unbound fraction (UF) wash (W) step, and purified products (1-4) following elution with maltose. B is protein left on the beads. The arrow denotes MBP-GGDEF. (SM is the size marker).

GTP and c-di-GMP standards (Fig. 9a). The positive control, Slr1143, revealed two peaks; one peak at 9 min whose identity is unknown, and the peak at 17 min corresponding c-di-GMP (Fig. 9b). The MBP-GGDEF chromatogram, however, had one peak at 15 min corresponding to GTP (Fig. 9c). Since no c-di-GMP peak was produced, MBP-GGDEF is not a diguanylate cyclase under these conditions. The lack of diguanylate cyclase activity *in vitro* led me to think that the GGDEF may play a regulatory role in MbaA phosphodiesterase activity.

The EAL domain is functional independent of the GGDEF domain

Tandem GGDEF-EAL proteins are not uncommon proteins within bacterial species. *Vibrio cholerae* has 10 tandem GGDEF-EAL proteins out of the 62 c-di-GMP-associated proteins in its genome. Several proteins with tandem GGDEF-EAL proteins within other bacterial

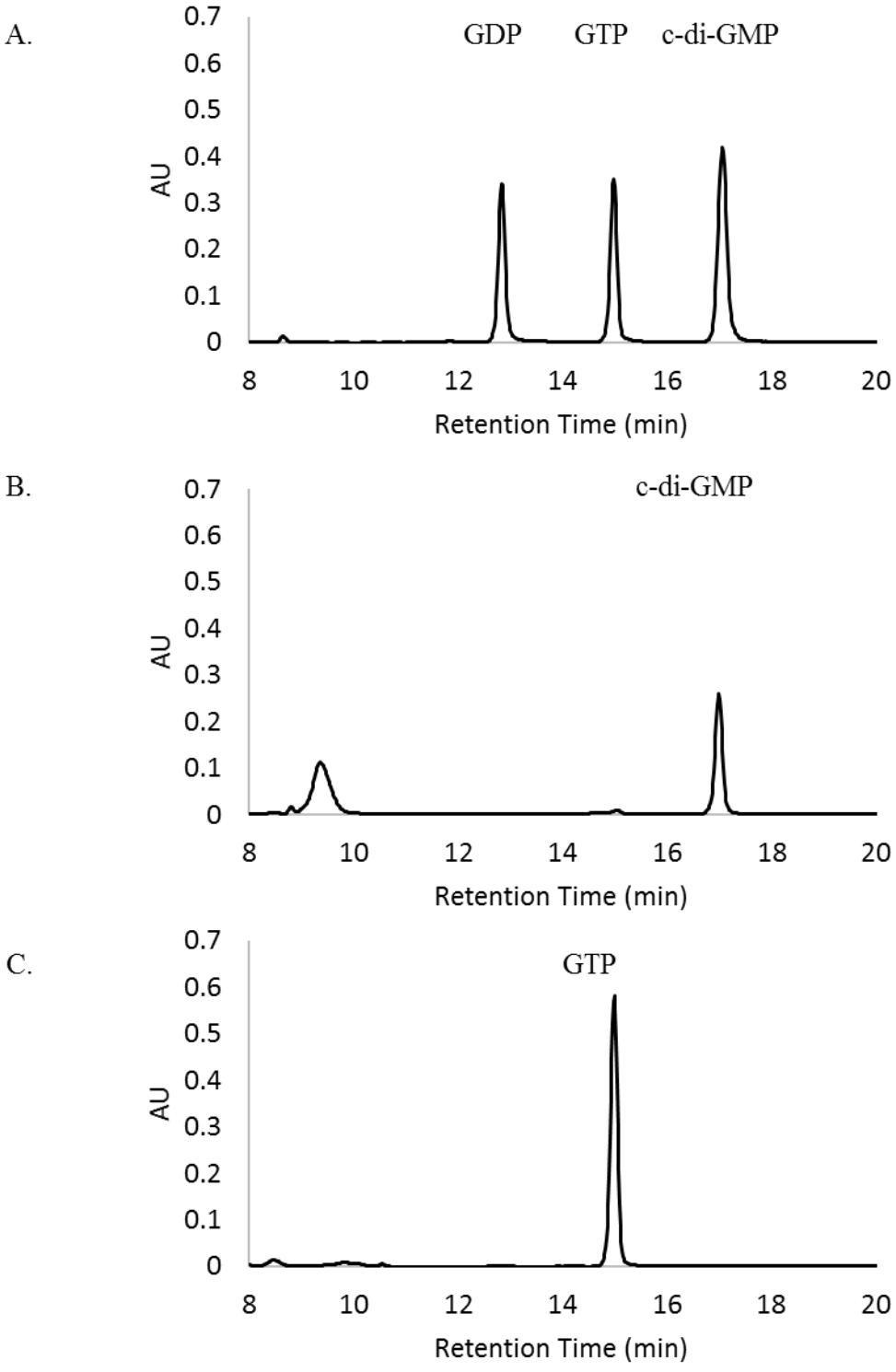

Fig 9. Chromatograms of (a) GDP, GTP, and C-di-GMP standards and the products of diguanylate cyclases assays for (b) Slr1143 and (c) MBP-GGDEF. (a) 100 μM GDP, GTP, and c-di-GMP standards. (b Slr1143; c MBP-GGDEF) Diguanylate cyclase assays were run with 2.5 μM enzyme and 100 mM GTP.

species have been characterized (Christen *et al.*, 2005, Kazmierczak *et al.*, 2006, Kuchma *et al.*, 2007). In those proteins that have an active phosphodiesterase activity but no active diguanylate cyclase activity, the GGDEF domain was shown to have a regulatory role (Christen *et al.*, 2005). So, my next hypothesis was that the GGDEF domain has a regulatory role within MbaA. For example, it may be possible that the GGDEF domain provides stability or another function that is necessary for phosphodiesterase activity to occur. If this were true, the EAL domain without the GGDEF domain would be nonfunctional and not be unable to linearize c-di-GMP. In order to establish if the GGDEF domain had a role in phosphodiesterase activity I produced a fusion protein of MBP-EAL and assayed its activity.

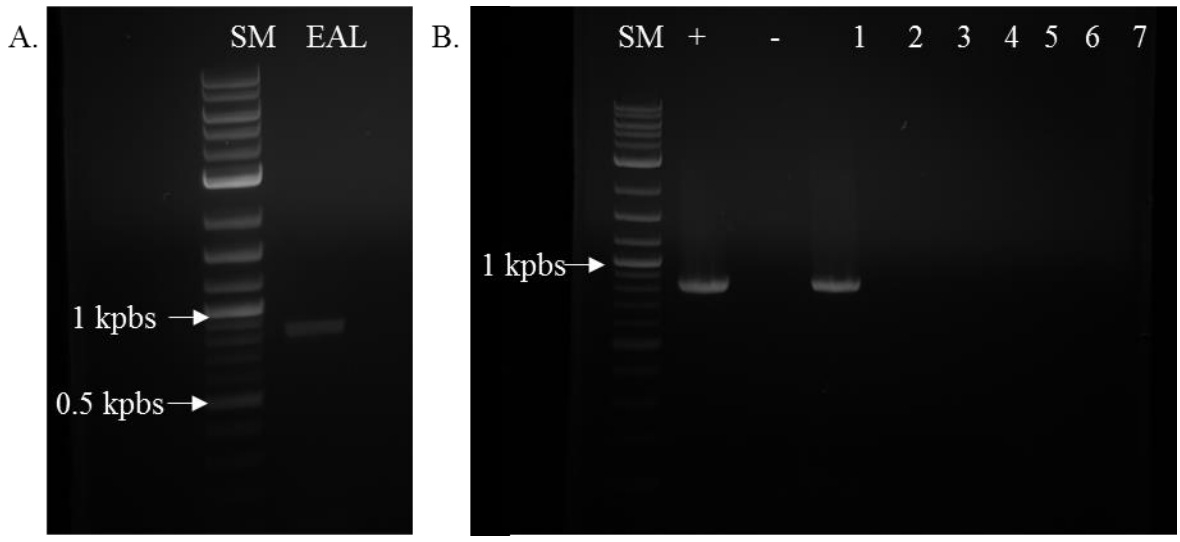

Fig 10. Construction of pWHIT4. (a) Amplification of the EAL domain from chromosomal DNA. The GGDEF amplicon was approximately 0.5 kbps based on size marker (SM). (b) Colony PCR of pWHIT4 within NEB express.

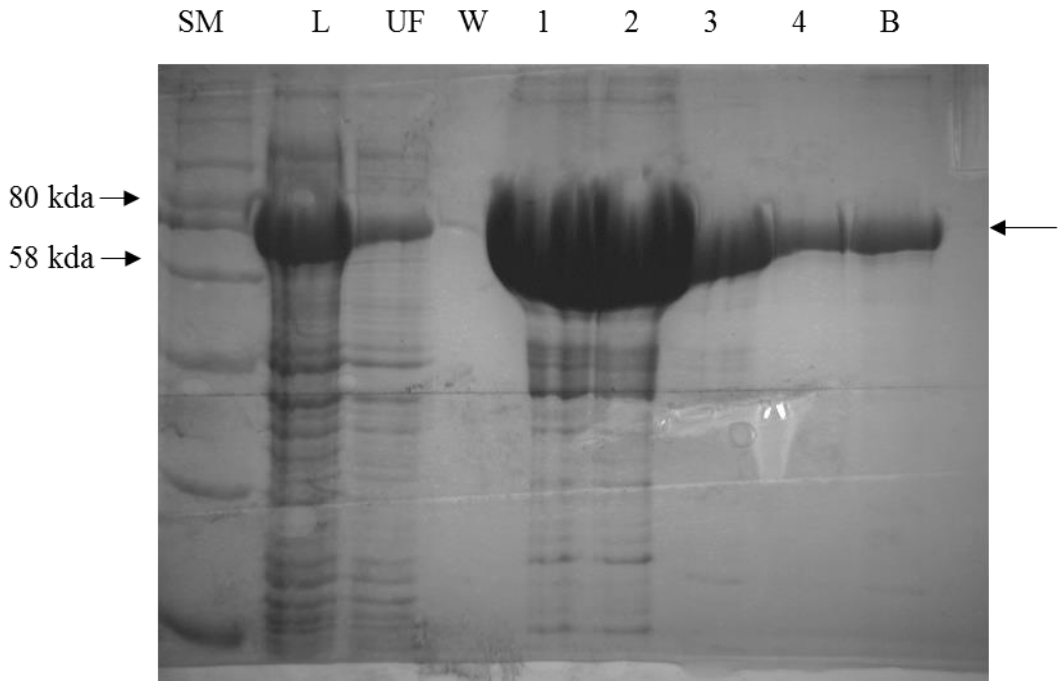

Fig 11. Purification of MBP-GGDEF. Lysate (L) from IPTG induced MBP-EAL in NEB Express was collected before binding with maltose resin. The unbound fraction (UF) wash (W) step, and purified products (1-4) following elution with maltose. B is protein left on the beads. The arrow denotes MBP-EAL. (SM is the size marker).

The sequence encoding the EAL domain of MbaA was amplified from chromosomal DNA to produce a PCR product of approximately 0.8 kbp (Fig. 10a). This product was cloned to produce pWHIT2 and confirmed by sequencing. The insert was then cloned into pMAL-c5x to produce pWHIT3. This plasmid was transformed into *E. coli* DH5α and then again into NEB express for maximal expression. The plasmid was confirmed in NEB express by colony

PCR to identify an amplicon of 0.8 kbp (Fig. 10b). The plasmid pWHIT2 was used as a positive control and successfully identified a 0.8 kbp amplicon. pMAL-c5x was used as a negative control and no amplicon was seen (Fig. 10b). Colonies 1-5 were screened; only colony 1 had an amplicon of 0.8 kbp (Fig. 10b).

Protein expression and purification was done with MBP-EAL and MBP-Cterm as described in the methods. MBP-Cterm is the fusion protein of maltose binding protein joined with the MbaA C-terminus containing both GGDEF and EAL domains. This fusion protein has been previously used to characterize MbaA as a phosphodiesterase and was used as a positive control in this study (Cockerell *et al.*, 2014). MBP-EAL was approximately 80 kDa (Fig. 11). The peaks in the chromatograms were then compared to standard peaks of 5' pGpG and c-di-GMP. The positive control, MBP-Cterm, had two peaks at 15 min and 17.5 min, which correspond to 5'pGpG and c-di-GMP, respectively (Fig. 12a). The HPLC chromatogram of MBP-EAL, however, had only one peak at 15 min, having only 5' pGpG detectable (Fig. 12b). MBP-EAL had a 100% conversion of c-di-GMP to 5' pGpG. MBP-Cterm in contrast had only 20% conversion of 5'pGpG from c-di-GMP. Within the same time frame, MBP-EAL had a higher output of 5' pGpG than MBP-Cterm indicating that MBP-EAL has a higher rate of product formation.

Because MBP-EAL seemed to produce more 5'pGpG in the same time frame, an assay was done to compare reaction rates. MBP-EAL and MBP-Cterm were incubated for 0, 30, 60, and 120 min and products were assayed as described above (Fig. 13). By 120 min MBP-EAL had converted all 100 µM of c-di-GMP to 5' pGpG, while MBP-Cterm had only converted roughly 20% of the c-di-GMP to 5' pGpG (Fig. 13). This time course confirmed that MBP-EAL produced 5'pGpG in a faster rate than MBP-Cterm. This indicates that the GGDEF domain is inhibitory to phosphodiesterase activity.

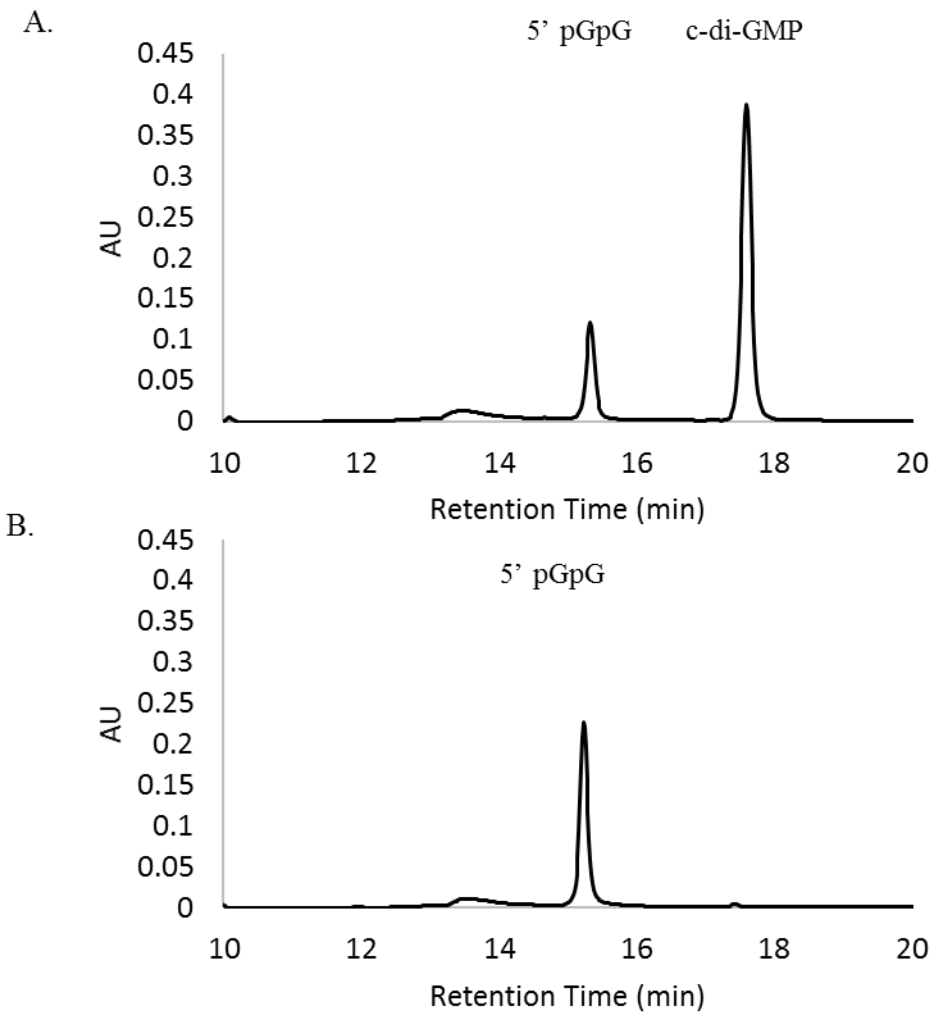

Fig 12. Chromatograms of phosphodiesterase assay products of (a) MBP-Cterm and (b) MBP-EAL. Phosphodiesterase assays were run with 2.0 μM or 2.5 μM enzyme for Cterm and EAL, respectively, and 100 μM c-di-GMP.

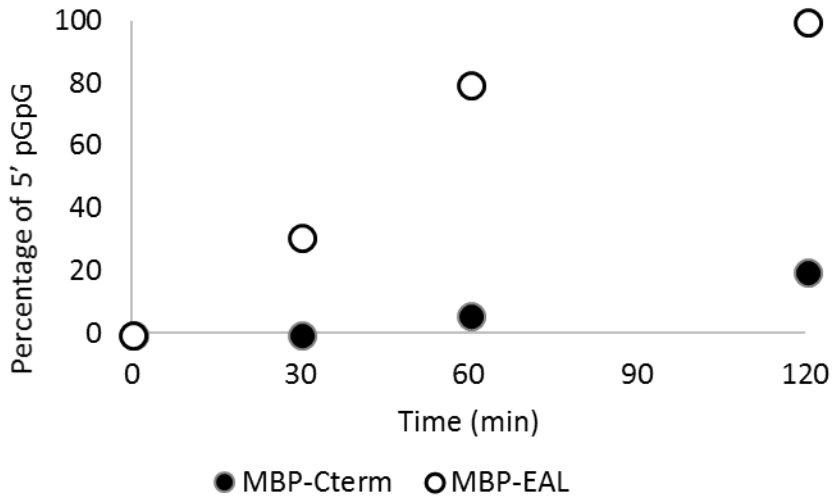

Fig 13. Time course shown percent 5' pGpG produced by MBP-Cterm and MBP-EAL. Phosphodiesterase assays were run with 2.0 μM and 2.5 μM enzyme for Cterm **(●)** and EAL **(○)**, respectively, and 100 μM c-di-GMP. Reactions were incubated for 0, 30, 60, and 120 minutes. The percent 5' pGpG produced vs time.

Role of GGDEF and EAL domains of MbaA *in vivo*

C-di-GMP concentration within the cell is highly regulated by different proteins and factors (Conner *et al.*, 2017). The *in vitro* data describe above have been collected in the absence of these native regulation systems. *In vitro*, the MbaA GGDEF domain has no activity and is inhibitory to the EAL domain. To determine whether *in vitro* results could be replicated *in vivo*, I performed domain deletions within the chromosome of *V. cholerae*. These mutations were designed to delete the various domains by homologous recombination. The GGDEF domain was deleted to see if its absence would disrupt MbaA phosphodiesterase activity *in vitro*. The EAL domain was deleted in order to see if the GGDEF domain may play a role in biofilm formation. A biofilm assay was then used to determine activity of the domains. Wildtype and Δ*mbaA* were used as comparisons for the mutants (Fig. 14). Biofilm formation exhibited by *ΔmbaA*, MbaAΔGGDEF, and MbaAΔEAL strains were significantly higher than the biofilm formation of wild-type MbaA. However,

biofilm formation exhibited by the MbaAΔGGDEF and MbaAΔEAL strains were not significantly different than biofilm formation by the *ΔmbaA* strain. These results were not expected as *in vitro* analysis did not show that the GGDEF domain was necessary for phosphodiesterase activity. The expected result would be that the GGDEF domain deletion would have diminished biofilms as phosphodiesterase activity would be increased (Fig. 14b). The high biofilms seen in the MbaAΔEAL strain were expected as without the EAL domain's phosphodiesterase activity MbaA would be unable to deplete the c-di-GMP concentration in the cell (Fig. 14b).

The deletion of the GGDEF domain did not support my *in vitro* results as biofilm formation was the same as the *mbaA* deletion strain. This could suggest that native regulatory elements play a part in biofilm formation that was not testable *in vitro*. However, it could also be that the domain deletions *in vivo* made the mutant protein unstable which hindered its expression. To test this, a Western Blot was performed to confirm proper expression of MbaA. I was unable to see expression, indicating that native levels of MbaA are too low to be analyzed using this method (Bond, unpublished results). To be able to detect MbaA, an overexpression plasmid, pMbaA, was used to generate the domain deletions as described in the Materials and Methods (Massie *et al.*, 2012). These modified plasmids were then transformed into the Δ*mbaA* background to produce the strains, Δ*mbaA* with pMbaAΔEAL and Δ*mbaA* with pMbaAΔGGDEF. pMbaA has a pEVS143 backbone so pEVS143 was used as a vector control (Massie *et al.*, 2012). Biofilm assays were conducted on Δ*mbaA* with pEVS143, Δ*mbaA* with pMbaA, Δ*mbaA* with pMbaAΔEAL, and Δ*mbaA* with pMbaAΔGGDEF (Fig. 15). The biofilm formed from strain Δ*mbaA* with pEVS143 was not significantly different than Δ*mbaA* pMbaAΔEAL, though it was significantly different than

pMbaAΔGGDEF (p-value 5.8 E^{-10}). Δ*mbaA* pMbaA was significantly different than Δ*mbaA* pMbaAΔEAL (p-value 3.8 E^{-16}), but not significantly different than Δ*mbaA* pMbaAΔGGDEF (Fig. 15). The increase in biofilm formation in Δ*mbaA* pMbaAΔEAL is due to the lack of phosphodiesterase and was seen in chromosomal domain deletions as well. The Δ*mbaA* pMbaAΔGGDEF strain had a decrease in biofilm comparable to the overexpression strain indicating that the GGDEF domain is not necessary for phosphodiesterase activity.

Plasmid based domain deletions were produced in order to confirm stability of these proteins. Western blots are currently under way to confirm that these domain deletion proteins are being expressed and that they are being expressed comparatively to full length MbaA encoded in pMbaA.

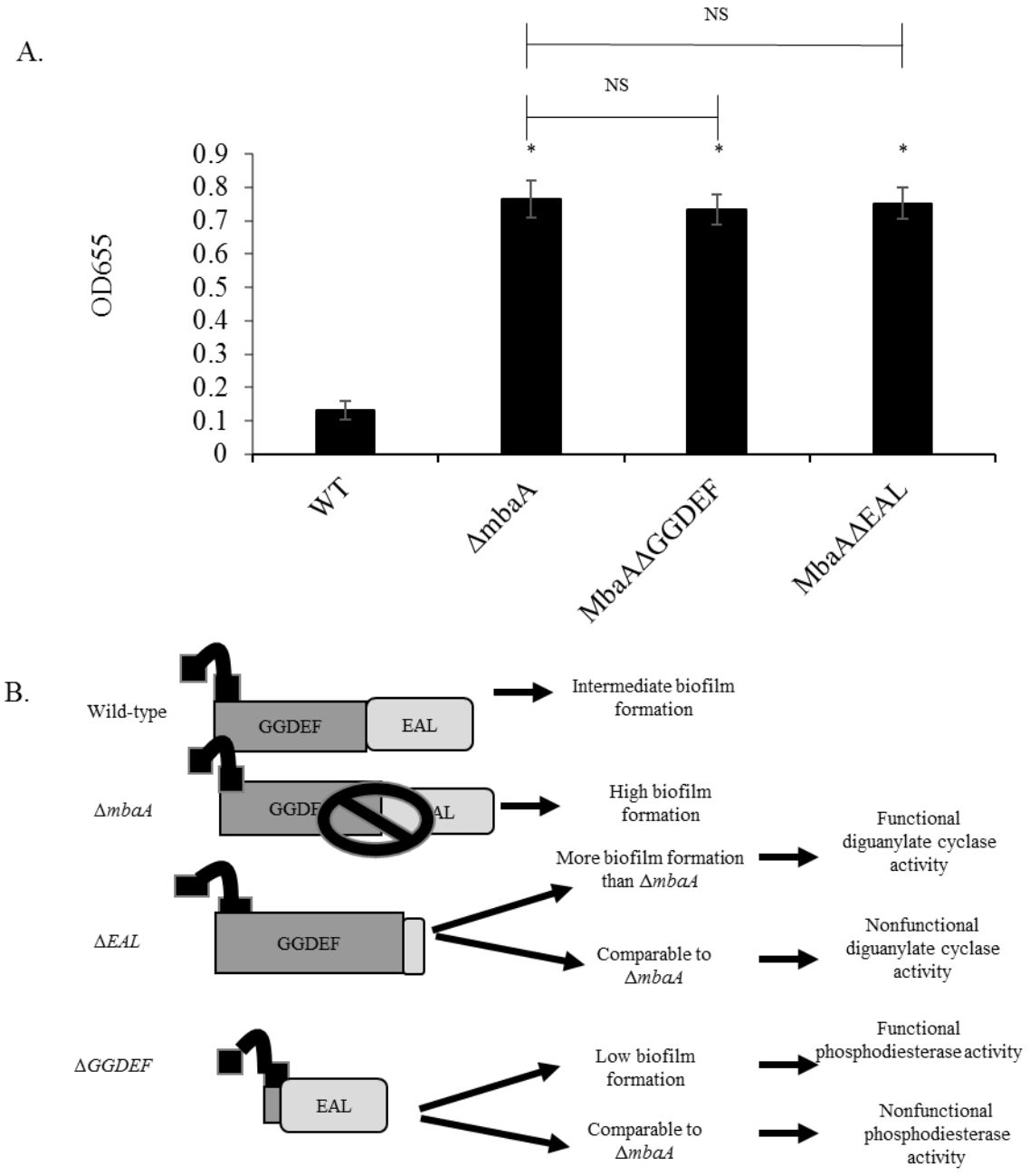

Fig 14. Biofilm formation in chromosomal domain deletions. (a) Biofilms were formed in borosilicate tubes in LB broth for 18 h at 27°C and quantified as described in Materials and Methods. Error bars show standard deviations of three biological replicates. A star indicates a significant difference between wild-type and mutants. Bars are a comparison between *ΔmbaA* and MbaAΔGGDEF and MbaAΔEAL which was not significant. A p-value of ≤ 0.05 was considered significant. WT, wild type. (b) Expected results of the biofilms performed.

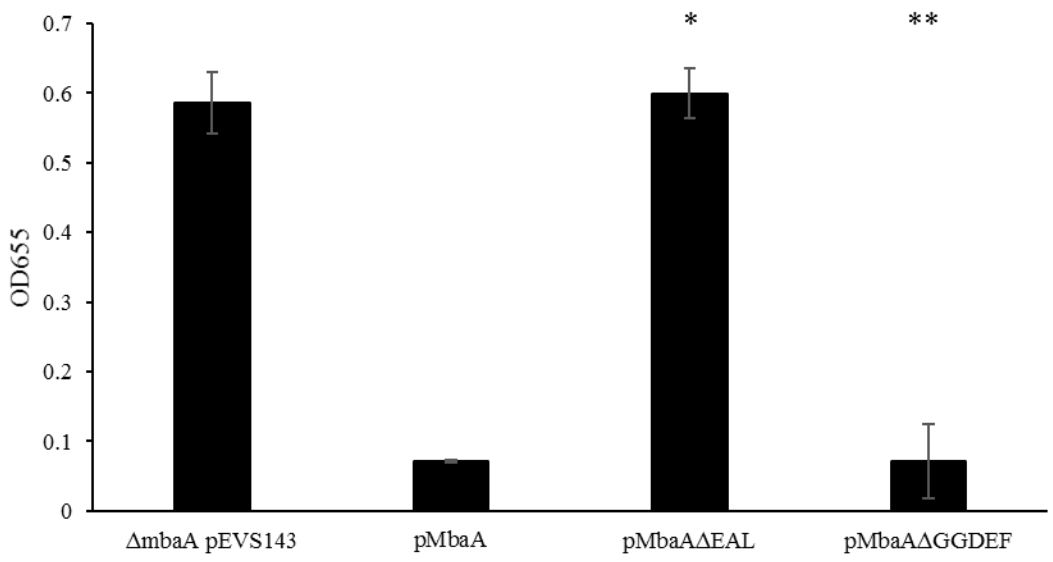

Fig 15. Biofilm formation plasmid based domain deletions. Biofilms were formed in borosilicate tubes in LB broth for 18 h at 37°C and quantified as described in Materials and Methods. Error bars show standard deviation of three biological replicates. One star indicates a statistical difference between *ΔmbaA* pEVS143 and *ΔmbaA* pMbaAΔGGDEF. Two stars indicate a statistical difference between *ΔmbaA* pMbaA and *ΔmbaA* pMbaAΔEAL. A p-value of ≤ 0.05 was considered significant. WT, wild type.

Discussion

V. cholerae is able to switch between a biofilm state and a motile, planktonic state in accordance with environmental signals. These environmental signals are then transduced into the cytoplasm where c-di-GMP concentration is controlled. An abundance of c-di-GMP enhances biofilm formation, while its degradation enhances motility. The NspS-MbaA signaling system is hypothesized to control c-di-GMP levels by sensing environmental polyamines. NspS binds to three polyamines: spermine, spermidine, and norspermidine. When NspS binds to spermine and spermidine it is hypothesized to cause MbaA to increase phosphodiesterase activity, decreasing c-di-GMP in the cell. When NspS binds to norspermidine, it is hypothesized that NspS inhibits MbaA phosphodiesterase activity within the cell, leading to an accumulation of c-di-GMP. MbaA is a tandem GGDEF-EAL protein. While the phosphodiesterase activity of MbaA has been linked with its EAL domain, my research set out to determine the function of the GGDEF domain (Cockerell *et al.*, 2014). I hypothesized that the GGDEF domain acted as a diguanylate cyclase when phosphodiesterase activity was inhibited.

MbaA has the potential to be a dual functional diguanylate cyclase phosphodiesterase

To first determine if this dual activity was possible, I performed an amino acid sequence alignment of the GGDEF-EAL domains between MbaA, DcpA, and ScrC. DcpA and ScrC are both dual functional diguanylate cyclase phosphodiesterases. I determined MbaA had the same catalytic amino acids necessary for both activities, as well as if MbaA

had the structural flexibility to accommodate both activities. The MbaA amino acid sequence retains all the catalytic amino acids necessary for both phosphodiesterase activity and diguanylate cyclase activity. MbaA also shares a similar number of amino acids between the two domains as ScrC and DcpA, indicating that the flexibility of the protein should be able to accommodate both enzymatic activities. However, MbaA also has a seven amino acid insertion seven amino acids upstream of the GGDEF motif. This insertion was not seen in either of ScrC and DcpA. How it affects the activity or protein structure of MbaA is not known.

MbaA does not have diguanylate cyclase activity *in vitro*

Because the MbaA had the potential for diguanylate cyclase activity I set out to determine if the GGDEF domain was a diguanylate cyclase *in vitro*. I did this by designing a fusion protein within the pMAL-c5x system. I cloned the GGDEF domain in frame with a maltose binding protein to produce MBP-GGDEF. This allowed me to be able to purify the MBP-GGDEF protein and perform a diguanylate cyclase assay with the fusion protein. I also purified a characterized diguanylate cyclase, Slr1143, and used it as a positive control in the same assay (Ryjenkov *et al.*, 2005). Slr1143 was able to convert all the GTP into c-di-GMP within 5 hours. MBP-GGDEF, however, did not produce any products after a 5-hour incubation with GTP at 37 °C. Diguanylate cyclases form c-di-GMP by binding to GTP and then dimerizing in an antiparallel fashion (Chan *et al.*, 2004, Ryjenkov *et al.*, 2005). It is possible that the MBP fusion to the GGDEF domain may be inhibiting dimerization. Cleaving the maltose binding protein from the MBP-GGDEF protein is possible; however, maltose binding protein can provide increased solubility to the protein so cleaving the MBP-GGDEF protein may cause the GGDEF domain to misfold and fall out of solution (Kapust &

Waugh, 1999). It is important to note that the positive control used for the diguanylate cyclase assay in this study, Slr1143, is also an MBP fusion. The possibility of the MBP preventing dimerization has already been analyzed. When first analyzing MBP diguanylate cyclase fusion proteins researchers added a coiled-coiled domain (CC), which enhance dimerization, to Slr1143 and analyzed formation of c-di-GMP. They found that c-di-GMP formation was not impacted by addition of CC domain indicating that the Slr1143 GGDEF domain alone was able to dimerize with MBP (Ryjenkov *et al.*, 2005). This data provides evidence that the lack of c-di-GMP formation was not due to an inhibition of dimerization but due to MBP-GGDEF not acting as a diguanylate cyclase in these conditions.

MbaA phosphodiesterase activity is independent of the GGDEF domain

Because the MBP-GGDEF protein was not an active diguanylate cyclase, I hypothesized that the GGDEF domain might play a role in the regulation of phosphodiesterase activity. To determine if the phosphodiesterase activity was affected by the GGDEF domain, I generated a MBP-EAL fusion protein. I also purified the previously characterized MBP-Cterm which is a fusion protein that has both the GGDEF and EAL domains of MbaA in frame with MBP (Cockerell *et al.*, 2014). I used MBP-Cterm as a positive control as well as a comparison for rate of phosphodiesterase activity. MBP-EAL protein converted 100 µM c-di-GMP completely into 5' pGpG, while MBP-Cterm was only able to convert 20 % of c-di-GMP to 5' pGpG in the same time frame. This indicated that the GGDEF domain is inhibitory. To confirm this, a time course assay was set up to determine how fast product formation took place with each of the two proteins. Reactions were set up for 0, 30, 60, and 120 minutes. MBP-EAL domain steadily created 5' pGpG, with the first formation of 5' pGpG seen at 30 minutes and c-di-GMP completely converted into 5' pGpG

at 120 minutes. MBP-Cterm, on the other hand, did not start 5' pGpG formation until 60 minutes and had only 20 % converted at 120 minutes. This confirms the results of the single time point phosphodiesterase assays and indicates that the GGDEF domain has an inhibitory effect on the phosphodiesterase activity under these conditions.

How the GGDEF domain inhibits phosphodiesterase activity in MbaA is not completely known at this time. The GGDEF domain of MbaA could be controlling the rate of phosphodiesterase activity to slow c-di-GMP degradation to prevent stressful conditions. Degrading c-di-GMP too fast may cause *V. cholerae* to disperse out of the biofilm or remain in a planktonic state when environmental conditions would be detrimental to the bacteria. A similar hypothesis has been assessed for why diguanylate cyclases in *V. cholerae* have different rates of biofilm formation (Massie *et al.,* 2012).

There is also the possibility of the GGDEF domain of MbaA acting as an allosteric inhibitor. Allosteric activation has been seen in the phosphodiesterase CC3396, a tandem GGDEF-EAL protein in *C. crescentus.* This protein is similar to MbaA in that the GGDEF domain of CC3396 is very close in amino acid sequence to active diguanylate cyclases. CC3396 has the motif GEDEF, rather than the characterized GGDEF motif and is unable to produce c-di-GMP; however, it can still bind GTP. When GTP is bound to CC3396 it further activates phosphodiesterase activity (Christen *et al.*, 2005). A similar regulation may be in place for MbaA where the GGDEF domain binds to a molecule, which causes inhibition of phosphodiesterase activity. The MbaA GGDEF domain has all the necessary catalytic amino acids for GTP binding; therefore, it may be able to bind to GTP or a similar molecule to inhibit the phosphodiesterase activity. The seven amino acid insertion seen in the GGDEF domain of MbaA may also modify the protein structure to bind to other molecules such as

GTP, ATP, or c-di-GMP. More research will have to be done if the GGDEF domain of MbaA is able to bind other molecules.

In vitro analysis is optimal for analyzing catalytic activity because it enables a clear analysis of activity; however, when isolating the protein, factors within the cell that may affect protein activity are also removed. To see if my *in vitro* data would remain consistent when MbaA activity was analyzed in its native system, *in vivo* domain deletions were constructed in the chromosome by homologous recombination. Since MbaA is a biofilm regulator, the effects of the GGDEF and EAL domain deletion within MbaA were analyzed using biofilm assays. Both the MbaAΔGGDEF and MbaAΔEAL strains showed increased biofilm formation similar to a Δ*mbaA* mutant. This was not as expected as I had hypothesized that MbaAΔGGDEF would have lower biofilm than wild-type since the GGDEF domain was inhibitory *in vitro*. The increase in biofilm seen in the MbaAΔGGDEF mutants indicates that within a native system MbaA phosphodiesterase activity requires the GGDEF domain. However, it also could be that the domain deletions affected the stability of the protein which caused the protein to be degraded. I tried to ensure that *mbaA* mutants were being expressed through western blot analysis; however, I was not able to detect native levels of MbaA in the cell. To overcome this issue, an overexpression plasmid, pMbaA, was used. Domain deletions were constructed within pMbaA to produce pMbaAΔGGDEF and pMbaAΔEAL. Biofilm assays were then conducted with strains expressing these altered proteins. pMbaAΔEAL strain showed an increase in biofilm formation since without the EAL domain, MbaA is unable to break down c-di-GMP. pMbaAΔGGDEF had very little biofilm formation, which is similar to pMbaA background. Minimal biofilm formation is indicative to decreased c-di-GMP concentration which indicates that pMbaAΔGGDEF is a

functional phosphodiesterase. This result confirms that the GGDEF domain is not necessary for phosphodiesterase activity. I failed to support my hypothesis that the GGDEF domain is inhibitory as pMbaAΔGGDEF was not statistically different than pMbaA. It is possible that the overexpression of MbaA is causing there to be more MbaA to degrade c-di-GMP, causing an artificially low biofilm formation. Without chromosomal mutants the inhibition caused by GGDEF domain in MbaA cannot be confirmed.

Future Directions

In this work, I have showed that the GGDEF domain is not necessary for phosphodiesterase activity of MbaA and that it has an inhibitory effect on phosphodiesterase activity *in vitro*. However, how and why the GGDEF domain inhibits phosphodiesterase activity still needs to be determined. A number of experiments can be done as the next steps to characterizing the domains of MbaA. First, point mutations within the GGDEF motif may shed some light into how the GGDEF domain is inhibiting the phosphodiesterase activity. The most common point mutation in GGDEF proteins is a glutamate to alanine substitution to produce a GGDAF motif. This abolishes the coordination of the magnesium ion and prevents diguanylate cyclase activity. However, MbaA does not have diguanylate cyclase activity; therefore, it is unclear if alteration of the glutamate to alanine in the GGDEF motif would have an effect on the phosphodiesterase activity as coordination of the magnesium ion is unlikely to have an effect as the binding of the domain to an allosteric inhibitor. If this mutation does not increase phosphodiesterase activity, other molecules, such as GTP, ATP, and c-di-GMP, would have to be substituted to examine if they are necessary for binding to an allosteric inhibitor. If the GGDEF domain is binding to an allosteric inhibitor, point mutations may be able to cause the same increase in phosphodiesterase activity as seen in the

in vitro assay of MBP-EAL. These point mutations would have to be constructed on the chromosomal copy of *mbaA* since the overexpression of MbaA prevented me from seeing a difference in biofilm formation of *ΔmbaA* with pMbaA and *ΔmbaA* with pMbaAΔGGDEF. If point mutations on the chromosome show a decrease in biofilm formation, indicative of increased phosphodiesterase activity, the next step would be to analyze what molecule is binding to the GGDEF domain of MbaA.

Molecules, such as GTP or ATP, could be added to the MBP-Cterm phosphodiesterase assay to see if they can replicate the phosphodiesterase rate seen in MBP-EAL. GTP was found to activate phosphodiesterase activity of CC3396 in *C. crescentus* but may inhibit MbaA phosphodiesterase activity. There is also a possibility the GGDEF domain of MbaA may bind to ATP as GGDEF domains are homologs to adenyl cyclases (Pei & Grishin, 2001). A modified GGDEF domain in YybT of *Bacillus subtilis* was found to be an ATPase, so GGDEF domains have the potential of binding to ATP (Rao *et al.*, 2010). Other molecules, such as other nucleotides, may be allosteric inhibitors as well, but further *in vitro* testing would have to be done to investigate this possibility. After the allosteric inhibitor is identified then point mutations from *in vivo* analysis could be applied to MBP-Cterm to see if these prevent binding to the inhibitor.

There is also the potential that the GGDEF domain plays a role in the NspS-MbaA signaling system. The GGDEF domain may provide flexibility to MbaA such that when NspS interacts with MbaA it could change the protein confirmation, causing the GGDEF domain to prevent the EAL domain from being able to cleave c-di-GMP. The Karatan lab is currently working on a method to analyze the NspS-MbaA interaction *in vitro*. MbaA has two hydrophobic regions that prevent its easy purification; this is combatted by using a

nanodisc which replicates the hydrophobic lipid bilayer and provides a surface for MbaA docking (Ritchie *et al.*, 2009). NspS can then be added to the nanodisc-MbaA complex and the change in activity can be assessed. If NspS inhibits phosphodiesterase activity, a GGDEF domain deletion can then be produced. If NspS requires the GGDEF domain to elicit a response, NspS will no longer be able to inhibit phosphodiesterase activity.

A confounding factor in this research was that while MBP-GGDEF was not able to produce c-di-GMP *in vitro*, pWHIT3 was unable to be expressed in NEB express. Overexpression of diguanylate cyclases have been noted as toxic to some *E. coli* strains (Ryjenkov *et al.*, 2005). DcpA and ScrC were characterized as dual-functional diguanylate cyclase phosphodiesterases *in vivo*, but it is unclear if they can function as diguanylate cyclases *in vitro* (Ferreira et al., 2008, Feirer et al., 2015). If MbaA is a functional diguanylate cyclase, then *ΔmbaA* with the pMbaAΔEAL should have produced a more biofilm formation than *ΔmbaA* with the empty vector. There is a possibility that at the time the biofilms were scored, biofilm formation was maximal and that there could be no more cells incorporated into the biofilm. In order to see if maybe pMbaAΔEAL produces a biofilm faster than *ΔmbaA*, a time course biofilm assay would have to be done. This would determine if pMbaAΔEAL and *ΔmbaA* formed biofilms at a similar rate. If pMbaAΔEAL forms biofilm at a faster rate than *ΔmbaA*, it would be indicative of diguanylate cyclase activity.

In order to determine if MbaA has diguanylate cyclase activity, c-di-GMP concentration can also be measured in a heterologous host. *E. coli* was used to characterize DcpA by overexpressing point mutants in the GGDEF and EAL domains. Glutamate to alanine substitutions were made in the EAL domain, to generate AAL, and in the GGDEF domain, to generate GGDAF. These substitutions abolish activity of phosphodiesterases and

diguanylate cyclases, respectively. When the AAL DcpA was overexpressed c-di-GMP concentration was high in *E. coli* while c-di-GMP concentration was low in the GGDAF DcpA overexpression strain. These results indicated that DcpA has both diguanylate cyclase and phosphodiesterase activity (Feirer *et al.*, 2015). *E. coli* is a good heterologous host because it naturally has low c-di-GMP concentration and it has none of the native regulation systems found in *V. cholerae*. In order to see if MbaA is a diguanylate cyclase c-di-GMP concentration can be calculated with MBP-GGDEF being expressed in DH5α. If c-di-GMP concentration is higher in this strain it would be indicative of an active diguanylate cyclase and would explain why we had issues transforming pWHIT3 into NEBexpress.

Conclusion

In this study, I set out to determine what role the GGDEF domain plays in MbaA. I found that while the GGDEF domain is not a functional diguanylate cyclase *in vitro* it plays a role in regulating phosphodiesterase activity of MbaA. The GGDEF domain inhibits phosphodiesterase activity which would decrease the rate that biofilm formation occurs in *V. cholerae*. Biofilm formation is believed to be important for *V. cholerae* survival within the environment and could also help the bacterial infection cycle. Because of how important biofilms are to the survival of *V. cholerae* within the environment, proteins that regulate biofilm formation may become important drug targets for treating and preventing *V. cholerae* infections. Clues on allosteric regulation of MbaA may shed light to how *V. cholerae* responds to stressful conditions that require transitions between planktonic and biofilm states. My research may also extend beyond *V. cholerae* research as many other proteobacteria have genes that encode NspS-MbaA like protein pairs that have not been characterized. This

research could extend to understanding of how these other homologs and regulate c-di-GMP concentration within their respective species.

References

Bellini D, Caly DL, McCarthy Y, Bumann M, An SQ, Dow JM, Ryan RP & Walsh MA (2014) Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. *Mol Microbiol* **91**: 26-38.

Bellini D, Horrell S, Hutchin A, Phippen CW, Strange RW, Cai Y, Wagner A, Webb JS, Tews I & Walsh MA (2017) Dimerisation induced formation of the active site and the identification of three metal sites in EAL-phosphodiesterases. *Sci Rep* **7**: 42166.

Bomchil N, Watnick P & Kolter R (2003) Identification and characterization of a *Vibrio cholerae* gene, *mbaA*, involved in maintenance of biofilm architecture. *J Bacteriol* **185**: 1384-1390.

Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U & Schirmer T (2004) Structural basis of activity and allosteric control of diguanylate cyclase. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 17084-17089.

Christen M, Christen B, Folcher M, Schauerte A & Jenal U (2005) Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *The Journal of biological chemistry* **280**: 30829-30837.

Cockerell SR, Rutkovsky AC, Zayner JP, Cooper RE, Porter LR, Pendergraft SS, Parker ZM, McGinnis MW & Karatan E (2014) *Vibrio cholerae* NspS, a homologue of ABC-type periplasmic solute binding proteins, facilitates transduction of polyamine signals independent of their transport. *Microbiol* **160**: 832-843.

Conner JG, Zamorano-Sanchez D, Park JH, Sondermann H & Yildiz FH (2017) The ins and outs of cyclic di-GMP signaling in *Vibrio cholerae*. *Current opinion in microbiology* **36**: 20-29.

Dey AK, Bhagat A & Chowdhury R (2013) Host cell contact induces expression of virulence factors and VieA, a cyclic di-GMP phosphodiesterase, in *Vibrio cholerae*. *J Bacteriol* **195**: 2004-2010.

Donlan RM & Costerton JW (2002) Biofilms: survival mechanisms of clinically relevant microorganisms. *Clinical microbiology reviews* **15**: 167-193.

Dunn AK, Millikan DS, Adin DM, Bose JL & Stabb EV (2006) New rfp- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and *lux* expression *in situ*. *Appl Environ Microbiol* **72**: 802-810.

Elasri MO & Miller RV (1999) Study of the response of a biofilm bacterial community to UV radiation. *Appl Environ Microbiol* **65**: 2025-2031.

Escobar LE, Ryan SJ, Stewart-Ibarra AM, Finkelstein JL, King CA, Qiao H & Polhemus ME (2015) A global map of suitability for coastal *Vibrio cholerae* under current and future climate conditions. *Acta tropica* **149**: 202-211.

Faruque SM, Albert MJ & Mekalanos JJ (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol Mol Biol Rev* **62**: 1301-1314.

Feirer N, Xu J, Allen KD, Koestler BJ, Bruger EL, Waters CM, White RH & Fuqua C (2015) A pterin-dependent signaling pathway regulates a dual-function diguanylate cyclase-phosphodiesterase controlling surface attachment in *Agrobacterium tumefaciens*. *mBio* **6**.

Ferreira RBR, Antunes LCM, Greenberg EP & McCarter LL (2008) *Vibrio parahaemolyticus* ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces. *J Bacteriol* **190**: 851-860.

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* **166**: 557-580.

Haugo AJ & Watnick PI (2002) *Vibrio cholerae* CytR is a repressor of biofilm development. *Mol Microbiol* **45**: 471-483.

Holmgren J, Lönnroth I, Månsson J & Svennerholm L (1975) Interaction of cholera toxin and membrane GM1 ganglioside of small intestine. *PNAS* **72**: 2520-2524.

Kapust RB & Waugh DS (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science : a publication of the Protein Society* **8**: 1668-1674.

Karatan E, Duncan TR & Watnick PI (2005) NspS, a predicted polyamine sensor, mediates activation of *Vibrio cholerae* biofilm formation by norspermidine. *J Bacteriol* **187**: 7434-7443.

Karatan E & Watnick P (2009) Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol Mol Biol Rev* **73**: 310-347.

Kazmierczak BI, Lebron MB & Murray TS (2006) Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* **60**: 1026-1043.

Kierek K & Watnick PI (2003) Environmental determinants of *Vibrio cholerae* biofilm development. *Appl Environ Microbiol* **69**: 5079-5088.

Kuchma SL, Brothers KM, Merritt JH, Liberati NT, Ausubel FM & O'Toole GA (2007) BifA, a cyclic-di-GMP phosphodiesterase, inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. *J Bacteriol* **189**: 8165-8178.

Lee SB, Park SK & Kim YS (2015) Production of bioactive chicken (*Gallus gallus*) follistatin-type proteins in *E. coli*. *AMB Express* **5**: 58.

Massie JP, Reynolds EL, Koestler BJ, Cong JP, Agostoni M & Waters CM (2012) Quantification of high-specificity cyclic diguanylate signaling. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 12746-12751.

McGinnis MW, Parker ZM, Walter NE, Rutkovsky AC, Cartaya-Marin C & Karatan E (2009) Spermidine regulates *Vibrio cholerae* biofilm formation via transport and signaling pathways. *FEMS microbiology letters* **299**: 166-174.

McNeill K & Hamilton IR (2003) Acid tolerance response of biofilm cells of *Streptococcus mutans*. *FEMS microbiology letters* **221**: 25-30.

Messing J (1983) New M13 vectors for cloning. *Methods in enzymology* **101**: 20-78.

Metcalf WW, Jiang W, Daniels LL, Kim SK, Haldimann A & Wanner BL (1996) Conditionally replicative and conjugative plasmids carrying *lacZ alpha* for cloning, mutagenesis, and allele replacement in bacteria. *Plasmid* **35**: 1-13.

Moorthy S & Watnick PI (2005) Identification of novel stage-specific genetic requirements through whole genome transcription profiling of *Vibrio cholerae* biofilm development. *Mol Microbiol* **57**: 1623-1635.

Pei J & Grishin NV (2001) GGDEF domain is homologous to adenylyl cyclase. *Proteins* **42**: 210-216.

Pérez-Mendoza D, Coulthurst SJ, Humphris S, Campbell E, Welch M, Toth IK & Salmond GP (2011) A multi-repeat adhesin of the phytopathogen, *Pectobacterium atrosepticum*, is secreted by a Type I pathway and is subject to complex regulation involving a non-canonical diguanylate cyclase. *Mol Microbiol* **82**: 719-733.

Prigent-Combaret C, Vidal O, Dorel C & Lejeune P (1999) Abiotic surface sensing and biofilm-dependent regulation of gene expression in *Escherichia coli*. *J Bacteriol* **181**: 5993-6002.

Rao F, Yang Y, Qi Y & Liang Z-X (2008) Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J Bacteriol* **190**: 3622-3631.

Rao F, See RY, Zhang D, Toh DC, Ji Q & Liang ZX (2010) YybT is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. *The Journal of biological chemistry* **285**: 473-482.

Reidl J & Klose KE (2002) *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS microbiology reviews* **26**: 125-139.

Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM & Sligar SG (2009) Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology* **464**: 211-231.

Römling U, Galperin MY & Gomelsky M (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* **77**: 1-52.

Ryjenkov DA, Tarutina M, Moskvin OV & Gomelsky M (2005) Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187**: 1792-1798.

Sanchez J & Holmgren J (2011) Cholera toxin - a foe & a friend. *The Indian journal of medical research* **133**: 153-163.

Schmidt AJ, Ryjenkov DA & Gomelsky M (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol* **187**: 4774-4781.

Silva AJ & Benitez JA (2016) *Vibrio cholerae* biofilms and cholera pathogenesis. *PLoS Negl Trop Dis* **10**: e0004330.

Sobe RC, Bond WG, Wotanis CK, Zayner JP, Burriss MA, Fernandez N, Bruger EL, Waters CM, Neufeld HS & Karatan E (2017) Spermine inhibits *Vibrio cholerae* biofilm formation through the NspS-MbaA polyamine signaling system. *The Journal of biological chemistry* **292**: 17025-17036.

Srivastava D, Harris RC & Waters CM (2011) Integration of cyclic di-GMP and quorum sensing in the control of *vpsT* and *aphA* in *Vibrio cholerae*. *J Bacteriol* **193**: 6331-6341.

Tarutina M, Ryjenkov DA & Gomelsky M (2006) An unorthodox bacteriophytochrome from *Rhodobacter sphaeroides* involved in turnover of the second messenger c-di-GMP. *The Journal of biological chemistry* **281**: 34751-34758.

Tchigvintsev A, Xu X, Singer A, *et al.* (2010) Structural insight into the mechanism of c-di-GMP hydrolysis by EAL domain phosphodiesterases. *Journal of molecular biology* **402**: 524-538.

Teschler JK, Zamorano-Sanchez D, Utada AS, Warner CJ, Wong GC, Linington RG & Yildiz FH (2015) Living in the matrix: assembly and control of *Vibrio cholerae* biofilms. *Nature reviews Microbiology* **13**: 255-268.

Tischler AD & Camilli A (2005) Cyclic diguanylate regulates *Vibrio cholerae* virulence gene expression. *Infection and immunity* **73**: 5873-5882.

Vanden Broeck D, Horvath C & De Wolf MJ (2007) *Vibrio cholerae:* cholera toxin. *The international journal of biochemistry & cell biology* **39**: 1771-1775.

Waldor MK & Mekalanos JJ (1994) Emergence of a new cholera pandemic: molecular analysis of virulence determinants in *Vibrio cholerae* 0139 and development of a live vaccine prototype. *J Infect Dis* **170**: 278-283.

Wang Y-C, Chin K-H, Tu Z-L, He J, Jones CJ, Sanchez DZ, Yildiz FH, Galperin MY & Chou S-H (2016) Nucleotide binding by the widespread high-affinity cyclic di-GMP receptor MshEN domain. *Nat Commun* **7**: 12481.

Whiteley CG & Lee DJ (2015) Bacterial diguanylate cyclases: structure, function and mechanism in exopolysaccharide biofilm development. *Biotechnology advances* **33**: 124-141.

WHO (2018) Cholera. p.^pp. World Health Organization.

Zhu J & Mekalanos JJ (2003) Quorum sensing-dependent biofilms enhance colonization in *Vibrio cholerae*. *Developmental cell* **5**: 647-656.

Vita

Whitney Bond was born in Redfield, Iowa. She graduated from Topsail High School in May 2011. She graduated *Cum Laude* from University of North Carolina Wilmington in May 2016 with a Bachelor's of Science in Biology and a minor in Forensic Anthropology. She started the Biology Master's program in August 2016 in the laboratory of Dr. Ece Karatan. She now resides in Kent, Washington with her husband and dog. Whitney Bond currently works as a technician at University of Washington Seattle.